

(TM)

**MPSrch\_nn** n.a. - n.a. database search, using Smith-Waterman algorithm

Tabular output not generated.

## SUMMARIES

TABLE default  
Gap 6

**Searched:**

Listing first 45

EST-STS

**EST-STS-TWC**

Database:						
EST-STS-THREE						
194:STS1	195:STS2	196:STS3	197:STS4	198:STS5	199:STS6	
200:STS7	201:STS8	202:STS9	203:STS10	204:STS11	205:STS12	
206:gnest1	207:gnest2	208:gnest3	209:gnest4	210:gnest5		
211:gnest6	212:gnest7	213:gnest8	214:gnest9	215:gnest10		
216:gnest11	217:gnest12	218:gnest13	219:gnest14			
220:gnest15	221:gnest16	222:gnest17	223:gnest18	224:gnest19		
225:gnest20	226:gnest21	227:gnest22	228:gnest23	229:gnest24		
230:gnest25	231:gnest26	232:gnest27	233:gnest28	234:gnest29		
235:gnest30	236:gnest31	237:gnest32	238:gnest33	239:gnest34		
240:gnest35	241:gnest36	242:gnest37	243:gnest38	244:gnest39		
245:gnest40	246:gnest41	247:gnest42	248:gnest43	249:gnest44		
250:gnest45						
Statistics: Mean 11.516; Variance 2.174; scale 5.296						
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.						
SUMMARIES						
Result No.	Score	Query Match Length	ID	Description	Pred. No.	
C 1	348	21.3	428_151	T94579	ye31g905.s1 Homo sapie	0.00e+00
C 2	245	15.0	405_151	T94272	ye31g905.r1 Homo sapie	0.00e+00
C 3	218	13.3	241_250	H8G124	human sts SHGC-17364.	0.00e+00
C 4	218	13.3	241_250	G29124	human sts SHGC-17364.	0.00e+00
C 5	81	5.0	405_41	H43270	yp05a1z.r1 Homo sapie	1.04e-10
C 6	78	4.8	541_113	N75653	yy29c01.r1 Homo sapie	8.69e-99
C 7	78	4.8	541_207	N75653	yy29c01.r1 Homo sapie	8.69e-99
C 8	73	4.5	533_144	T55854	yc24b03.s1 Homo sapie	5.24e-88
C 9	73	4.5	533_200	G24226	human sts Wt-11415.	5.24e-88
C 10	71	4.3	311_79	N40107	ym72b09.r1 Homo sapie	3.96e-85
C 11	71	4.3	336_169	N51302	ma16f08.r1 Life Tech	3.96e-88
C 12	66	4.0	266_35	H24821	y142f01.r1 Homo sapie	1.59e-75
C 13	65	4.0	396_134	R55530	yy19c1z.r1 Homo sapie	1.28e-77
C 14	60	3.7	306_215	W10705	ma47g07.r1 Soares mou	3.39e-66
C 15	60	3.7	306_174	W10705	ma47g07.r1 Soares mou	3.39e-66
C 16	60	3.7	421_139	R74158	yy19g905.r1 Homo sapie	3.39e-66
C 17	60	3.7	438_76	N26985	ym72b09.s1 Homo sapie	3.39e-64
C 18	60	3.7	586_170	W53184	md1gao3.r1 Soares mou	3.39e-66
C 19	59	3.6	589_193	W97818	mg03c04.r1 Soares mou	2.48e-65
C 20	56	3.4	320_144	T66009	yc24b03.r1 Homo sapie	8.71e-55
C 21	56	3.4	299_151	T91693	ye25j05.r1 Homo sapie	3.79e-44
C 22	43	2.6	459_31	H10989	ym07a06.s1 Homo sapie	9.88e-34
C 23	41	2.5	461_56	H94102	yy56g11.r1 Homo sapie	2.30e-30
C 24	38	2.3	392_95	N64805	yz31c01.s1 Homo sapie	1.98e-25
C 25	36	2.2	335_47	R02212	C.elegans CDNA clone	1.17e-22
C 26	35	2.1	280_98	R02212	pK05a03.s1 Caenorhabd	1.17e-20
C 27	35	2.1	360_46	CEIK04F3F	C.elegans CDNA clone	1.17e-20
C 28	34	2.1	360_47	CEIK04F3F	C.elegans CDNA clone	4.17e-19
C 29	34	2.1	360_45	CEIK05IHF	C.elegans CDNA clone	4.17e-19
C 30	34	2.1	429_44	CAEKO09DF	C.elegans CDNA clone	4.17e-19
C 31	32	2.0	323_244	NNA67770	mm24c01.r1 Stratiagene	4.51e-16
C 32	32	2.0	323_244	NNA67770	mm24c01.r1 Stratiagene	4.51e-16
C 33	33	2.0	360_45	CEIK06O3F	C.elegans CDNA clone	1.41e-17
C 34	33	2.0	360_47	CEIK06O3F	C.elegans CDNA clone	1.41e-17
C 35	31	1.9	379_224	CEIC4671	C.elegans CDNA clone	1.36e-14
C 36	31	1.9	379_39	C08467	C.elegans CDNA clone	1.36e-14
C 37	28	1.7	280_231	NSA62630	zm02h01.r1 Stratiagene	2.57e-10
C 38	28	1.7	280_231	AAO54927	SMWFCA800SK Brugia ma	1.27e-07
C 39	26	1.6	331_11	AAO54927	SMWFCA800SK Brugia ma	1.27e-07
C 40	26	1.6	331_223	BWA54927	md	

44 25 1.5 447 109 R40271 yf80f04.s1 Homo sapie 2.48e-06  
45 24 1.5 450 216 W13185 ma81d02.r1 Soares mou 4.42e-05

ALIGNMENTS

RESULT 1  
LOCUS T94579 428 bp mRNA EST 24-MAR-1995  
DEFINITION yef1905.s1 Homo sapiens cDNA clone 119384 3'.  
ACCESSION T94579  
KEYWORDS 9728067  
SOURCE EST.  
human clone-119384 library-Stratagene lung (#937210)  
vector-pbluescript SK- host-SOLR cells (kanamycin resistant)  
primer--21ml3 Rsttel-EcoRI Rsttel2-XhoI Normal lung tissue from a 72  
year old male. Cloned unidirectionally. Primer: Oligo dt. Average  
insert size: 1.0 kb; Uni-ZAP XR Vector; 5' adaptor sequence:  
5'-GAATTCGGCAGAG-3'; 3' adaptor sequence:  
5'-CTCGAGTTTCTTTTCTTTT-3'.

ORGANISM  
Homo sapiens  
Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;  
Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE  
1 (bases 1 to 428)  
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,  
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Maira,M.,  
Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,  
Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.  
WashU-Merck EST Project  
Unpublished (1995)

TITLE  
JOURNAL  
COMMENT  
Contact: Wilson RK  
WashU-Merck EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu  
High quality sequence stops: 241  
Source: IMAGE Consortium, LNL  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.

FEATURES  
source  
Location/Qualifiers  
1..428  
/organism="Homo sapiens"  
/clone="119384"

BASE COUNT 112 a 103 c 124 g 76 t 13 others

ORIGIN  
Query Match 21.3%; Score 348; DB 151; Length 428;  
Best Local Similarity 94.4%; Pred. No. 0.00e+00;  
Matches 386; Conservative 0; Mismatches 18; Indels 5; Gaps 5;

Db 4 gaagaggcacaaacaaagattttttgcaagtgaaggaggcgc-aggagaccga 62  
|||||  
Cp 1624 GAAGAGGGGCAACAAACAAAGATTATTTTGCAGTAAGGAGGCCAGCAGAGACCA 1565  
|||||  
Db 63 gaaaaaagaagaagcgtgagagagagaagccttgataaagaagaccacaagaagaag 122  
|||||  
Cp 1564 GAAAAAAGAGAGCAAGGCTGAGAGCAGAAAGCCTGATTAAGAGACCAAGAAAGGC 1505  
|||||  
Db 123 ctgcagagcagatgtcgcccccagggaaaaacccagagagcagctgtagagtagc 182  
|||||  
Cp 1504 CTGCAGAGACCAAGATTGCGGCCGCCCAAGGAAACCCAGAGAGAGCTGTAGAGTAGTC 1445  
|||||  
Db 183 tgggcccagctcaaaagtgcagcttgagaggggtttaa-gaactcaattccaggtcagca 241  
|||||  
Cp 1444 TGGGCCCAACCTCAAGCTGGGAGCTGAGAGGCGCTTTAGGAGCTCAATTCAGAGTCAGCA 1385  
|||||  
Db 242 ttgcagagantgtctgaacacagcagccttgcggcagctt-ctgnaaacacmgcccgcn 300  
|||||  
Cp 1384 TTTTGCAGAGGATTGCTGAACAACAGGCGCTGTGCGGAGCTTGTCTGGAACAAGCCGCCGCC 1325  
|||||  
Db 301 tgcacaactgtagaagctgagaccgttcccgagagnttgtagatagagccatcagctt-cg 359  
|||||

|||||  
Cp 1324 TGCAACAGCTGTAGAGAGCTGAGACCGTTCCCGAGGATTTGGATAGAGCCCATAGCTTTGCC 1265  
|||||  
Db 360 ctggcagaagcttctctgtncagggtctngggcatgcttcaggttcaga 408  
|||||  
Cp 1264 CTGGCAGAACGTGTCTGTCTCCAGGCGCTGGGCCATGCT-CAGGTTCAGA 1217  
|||||

RESULT 2  
LOCUS T94272 405 bp mRNA EST 24-MAR-1995  
DEFINITION yef1905.r1 Homo sapiens cDNA clone 119384 5'.  
ACCESSION T94272  
KEYWORDS 9727760  
SOURCE EST.  
human clone-119384 library-Stratagene lung (#937210)  
vector-pbluescript SK- host-SOLR cells (kanamycin resistant)  
primer-M13RPI Rsttel-EcoRI Rsttel2-XhoI Normal lung tissue from a 72  
year old male. Cloned unidirectionally. Primer: Oligo dt. Average  
insert size: 1.0 kb; Uni-ZAP XR Vector; 5' adaptor sequence:  
5'-GAATTCGGCAGAG-3'; 3' adaptor sequence:  
5'-CTCGAGTTTCTTTTCTTTT-3'.

ORGANISM  
Homo sapiens  
Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;  
Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE  
1 (bases 1 to 405)  
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,  
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Maira,M.,  
Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,  
Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.  
WashU-Merck EST Project  
Unpublished (1995)

TITLE  
JOURNAL  
COMMENT  
Contact: Wilson RK  
WashU-Merck EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu  
High quality sequence stops: 214  
Source: IMAGE Consortium, LNL  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.

FEATURES  
source  
Location/Qualifiers  
1..405  
/organism="Homo sapiens"  
/clone="119384"

BASE COUNT 64 a 113 c 102 g 115 t 11 others

ORIGIN  
Query Match 15.0%; Score 245; DB 151; Length 405;  
Best Local Similarity 91.6%; Pred. No. 0.00e+00;  
Matches 371; Conservative 0; Mismatches 14; Indels 20; Gaps 20;

Db 1 tctgacctgagcatggcccccagcccttgacaagaacagcttctgcaggagaaactgat 60  
|||||  
Cp 1217 TCTGAACCTGACAGATGCCCCCAGCCCTTGACAAAGACAGTCTTGCCAAAGACTAT 1276  
|||||  
Db 61 gggcctatcccaatcctcgtgggaagcgtccagcttcaacagctgtgcagcgggcg-ctg 119  
|||||  
Cp 1277 GGGCTCTATCCCAATTCCTCGGGAAGGTCAGAGCTCTACAGCTGTGCAAGCGGGCGGCTG 1336  
|||||  
Db 120 ttccagcaaaagtgtccgacag-cctggtgttcagcaactcctgaaatgtctgcaactg 178  
|||||  
Cp 1337 TTCAGCAAAAGCTGCCACAGAGCGCTGTGTTACACCAACTCTGCATAATGTGTGCACTGG 1396  
|||||  
Db 179 gaattgagtcgttaagccccctcagctccagcttggagcttgaggccaggaattcactc 238  
|||||  
Cp 1397 -AATTGAGTCGTTAAAGCCCTTCAGTCCAGCTTTGAGCTGGGCGCAAG-AT-CACTC 1453  
|||||  
Db 239 ttacagccttctcctctctggtgttctcctctgggggcccgaatctggnctcttgcagc 298  
|||||  
Cp 1454 T-ACAGCCT-GCCT-CTGGG-TTTTCCCTGGGGG-CCGCAATCTG-CTCCT-GCAGGC 1506  
|||||



RESULT	5	HA3270	405 bp	mRNA	EST	31-JUL-1995
LOCUS		yp05a12.r1	Homo sapiens cDNA clone 186526 5'			similar to gb:M80927
DEFINITION		CARTILAGE GLYCOPROTEIN-39 PRECURSOR (HUMAN);				
ACCESSION		HA3270				
NID		9191322				
KEYWORDS		EST.				
SOURCE		human clone=186526 library=Soares breast 3Mdbst vector=pt773d (Pharmacia) with a modified polylinker host=DH10B (ampicillin resistant) primer=M13RP1 Rsite1-Not I Rsite2-Eco RI Adult human 1st strand cDNA was primed with a Not I - 0150(dt) primer [5' TGTTCACATCTCGAAGTGGAGCGGCCGCCCTTTTCTTTTCTTTTCTTTT 3'] double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT773 vector (Pharmacia). Library went through one round of normalization to a Cot = 20. Library constructed by Bento Soares and M.Fatima Bonaldo.				
ORGANISM		Homo sapiens				
REFERENCE		Eukaryotes; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Homalidae; Homo.				
AUTHORS		1 (bases 1 to 405) Hallier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,J., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,T., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaekis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.				
TITLE		The Washu-Merck EST Project				
JOURNAL		Unpublished (1995)				
COMMENT		Contact: Wilson RK Washu-Merck EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@wustl.wustl.edu High quality sequence stops: 323 Source: IMAGE Consortium, LLNL This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Location/Qualifiers 1..405 /organism="Homo sapiens" /clone="186526"				
BASE COUNT		106 a 112 c 95 g 86 t				6 others
ORIGIN						
Query Match		5.0%; Score 81; DB 41; Length 405;				
Best Local Similarity		63.3%; Pred. No. 1.04e-104;				
Matches		197; Conservative 0; Mismatches 113; Indels 1; Gaps 1;				
Db	1	gctgctccagatgcccttgacgcgttcctcctgacacacatcatctacagcttgcca	60			
Oy	120	gcttcctgcgcccaagagacttgagacccacagcctttggacacacactatctacgcttgccg	179			
Db	61	atataagcaagatcacatcgacacacctggagtggaatgagtgaagcgtctcagcgcatgc	120			
Oy	180	gcattaccacacccacagctgacacacactgactgagtggaatgacgagctctgtaccacgaggt	239			
Db	121	tcaaacactcaagaacaggaaccccaactcaaacctctctgctctgtagagagatga	180			
Oy	240	tcaattggccttgaaagatgaaatcccaagcttaaaagccctgtttaccatctggaagctgga	299			
Db	181	nonttgggtctcaagaatttccaaagatagccttcaaacacccagagctgcgcgacttca	240			
Oy	300	atttgggcacatcagaaatttcacagatattgtagccacggcccaacaacgttagacttttg	359			
Db	241	tcaagtcagtaacgcacattctcgagaccccatgagcttgattgagtgagctgagcttgc	300			
Oy	360	tcaactctgcgcacatcagaggttttcgccaatatacaggtttgacggc-ccttgacctgactgg	418			

DB	301	gtctaccctgg	311
QY	419	GAGTACCACGAG	429
RESULT	6	N75653	541 bp
LOCUS			MRNA
DEFINITION			EST
ACCESSION			02-APR-1996
NID			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
BASE COUNT			
ORIGIN			
Query Match	4.8%	Score 78;	DB 113; Length 541;
Best Local Similarity	66.7%	Pred. No. 8.6e-99;	
Matches 166;	Conservative 0;	Mismatches 82;	Indels 1; Gaps 1
Db	27	ctggcaggttagtgcgtcgtcgtcgtcctccagagagatcgtccatacaactggttg	86
QY	19	CTGGCAGGTTTCAATGTCCTCGTCGATGATGCCATGGGCGCTCTCTGCAAAACGTGCTG	78
Db	87	ctacttaccacactggtcccgagccggcagacacaggaana-ttcaacctggaatat	145
QY	79	CTACTTACCAACATGGGCCAGTACAGACAGAGGGAGGCTGCTCTGCTGCCAAGGACTT	138
Db	146	tgaccccttctatgctctcatcatcatatctatcttgccagatcgaataacaagaagt	205
QY	139	GGACCCAGCCCTTGGACCCACCTCATCTTACGCTTGCTGGCATGTACCAACCCACGACT	198
Db	206	tatcatcaaggaacaagatgaagtgaagctctaccagacacatcaacagctctcaaacaa	265



QY 199 GAGCACCAGCTGAGTGGATGACAGACTCTCTACACAGAGTTCAATGGCCTGAAGAAGAT 258  
 Db 266 gaatcccaa 274  
 QY 259 GAATCCCAA 267  
 RESULT 7 541 bp mRNA EST 02-APR-1996  
 LOCUS N75653  
 DEFINITION yv29c01.r1 Homo sapiens cDNA clone 244128 5' similar to gp:M80927  
 CARRIER GLYCOPROTEIN-39 PRECURSOR (HUMAN)..  
 ACCESSION N75653  
 NID 91238231  
 KEYWORDS EST.  
 SOURCE human clone=244128 primer=reverse ET library=Soares fetal liver  
 spleen INFES vector=PT73D (Pharmacia) with a modified polylinker  
 host=DH10B (ampicillin resistant) Rstrel-Pac I Rstrel2-Eco RI liver  
 and spleen from a 20 week-post conception male fetus. 1st strand  
 cDNA was primed with a Pac I - 0190(dT) primer [5].  
 AACTGGAAGATTAATTAAGATCTTTTCTTTTCTTTT 3'1, double-stranded  
 cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac  
 I and cloned into the Pac I and Eco RI sites of the modified pT73  
 vector. Library went through one round of normalization. Library  
 constructed by Bento Soares and M.Fatima Bonaldo.  
 ORGANISM Homo sapiens  
 Eukaryote; Metazoa; Eumetazoa; Bilateria; Coelomata;  
 Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;  
 Sarcopterygii; Chonata; Tetrabrata; Amniota; Mammalia; Theria;  
 Eutheria; Archonta; Primates; Catarrhini; Homidae; Homo.  
 1 (bases 1 to 541)  
 Hiller,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,  
 Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,  
 Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,  
 Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and  
 Wilson,R.  
 TITLE The Washu-Merk EST Project  
 JOURNAL Unpublished (1995)  
 COMMENT Contact: Wilson RK  
 Washu-Merk EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 High quality sequence stops: 441  
 Source: IMAGE Consortium, LNL  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.lnl.gov) for further information.  
 FEATURES NCBI g1: 1238231  
 source location/Qualifiers  
 1..541  
 /organism="Homo sapiens"  
 /clone="244128"  
 /note="human"  
 BASE COUNT 149 a 134 c 117 g 136 t 5 others  
 ORIGIN  
 mRNA  
 BASE COUNT 149 a 134 c 117 g 136 t 5 others  
 ORIGIN  
 Query Match 4.88; Score 78; DB 207; Length 541;  
 Best Local Similarity 66.7%; Pred. No. 8.69e-99;  
 Matches 166; Conservative 0; Mismatches 82; Indels 1; Gaps 1;  
 Db 27 ctggagaggtgtgtgtctgtctgtctccagagagatgcctcaactgtgttg 86  
 QY 19 ctggcaggggtttcatgtctctgatgatccatggggctgtgcgcaaaactggctg 78  
 Db 87 ctacttaccactgtgtccacagccgagaaacagaaanaa-ctcaaccctgagaatat 145  
 QY 79 ctacttcacacaaactggggccagtaacagacggggagggcctgccttctcccaagact 138

Db 146 tgacccttcctatgtctcatcatcattcatcgcgcacatcgaaacaaaggt 205  
 QY 139 GGACCCAGACCTTTGGACACCACTCATCTACGCTTCCTGGCATGACCAACACCAAGCT 198  
 Db 206 tatcatcaagacaagagtgtgtgtctctacacagacatcaacagtctcaaaccaa 265  
 QY 199 GAGCACCAGCTGAGTGGATGACAGACTCTCTACACAGAGTTCAATGGCCTGAAGAAGAT 258  
 Db 266 gaatcccaa 274  
 QY 259 GAATCCCAA 267  
 RESULT 8 533 bp mRNA EST 20-FEB-1995  
 LOCUS T65854  
 DEFINITION yv24b03.s1 Homo sapiens cDNA clone 81581 3' similar to SP:S27879  
 S27879 SECRETORY PROTEIN YN-1 PRECURSOR -.  
 ACCESSION T65854  
 NID 9674899  
 KEYWORDS EST.  
 SOURCE human clone=81581 library=Stratagene lung (#937210)  
 vector=PB1uescript SK- host=SOLR cells (kanamycin resistant)  
 primer=23m13 Rstrel-EcoRI Rstrel2-XhoI Normal lung tissue from a 72  
 year old male. Cloned unidirectionally. Primer: Oligo dT. Average  
 insert size: 1.0 kb; Uni-ZAP XR Vector; 5' adaptor sequence:  
 5'-GAATTCGGCAGAG-3'; 3' adaptor sequence:  
 5'-CTGAGTGTCTTTTCTTTTCTTTT-3'.  
 ORGANISM Homo sapiens  
 Eukaryote; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;  
 Eutheria; Primates; Catarrhini; Homidae; Homo.  
 1 (bases 1 to 533)  
 Hiller,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,  
 Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,  
 Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevasakis,E.,  
 Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.  
 TITLE Washu-Merk EST Project  
 JOURNAL Unpublished (1995)  
 COMMENT Contact: Wilson RK  
 Washu-Merk EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 High quality sequence stops: 307  
 Source: IMAGE Consortium, LNL  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.lnl.gov) for further information.  
 FEATURES NCBI g1: 1238231  
 source location/Qualifiers  
 1..533  
 /organism="Homo sapiens"  
 /clone="81581"  
 BASE COUNT 140 a 123 c 136 g 128 t 6 others  
 ORIGIN  
 Query Match 4.5%; Score 73; DB 144; Length 533;  
 Best Local Similarity 70.1%; Pred. No. 5.24e-89;  
 Matches 131; Conservative 0; Mismatches 55; Indels 1; Gaps 1;  
 Db 250 tggacaagaggtgaagagccctgtgtgtcagagatttccagtgaaagtcacatgtataa 309  
 Cp 1137 TGGATAGGGGGGTATGGGCTGTGTGCGAGAGACCGGCAAGTCATCTAAGTCCAGT 1078  
 Db 310 gaccagatcatgctctccacaggttaattctttaaagaactgaacctgtgtctcata 369  
 Cp 1077 GCCCAACATAGCCCCGCCCACTCCCTTCTGTAGATAGTGACCTTGTTGTTGAAG 1018  
 Db 370 ctcttcacatcatatgacccacacactgtgttcccttgactnctgaagggagacctgtg 429  
 Cp 1017 CTCTCCACATCAATCAAGCCACCACT-GGTTGTCGGGAGAGATGTAGGGACCTTCTG 959  
 Db 430 atctctgg 436

|||||  
CP 958 ATCTGTG 952

RESULT 9  
LOCUS G24226 533 bp DNA STS 31-MAY-1996  
DEFINITION human STS WI-14165.  
ACCESSION G24226  
NID g1344552  
KEYWORDS STS sequence; primer; sequence tagged site.  
SOURCE human STS derived from sequences in dbEST and the Unigene collection.

ORGANISM  
Homo sapiens  
Eukaryota; Eukaryotes; Metazoa; Chordata;  
Vertebrata; Mithertia; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 533)  
Hudson, T.  
Whitehead Institute/MIT Center for Genome Research; Physically  
Mapped STS  
Unpublished (1995)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Contact: Thomas Hudson  
Whitehead Institute/MIT Center for Genome Research  
Whitehead Institute for Biomedical Research  
9 Cambridge Center, Cambridge MA 02142 USA  
Tel: 617 252 1900  
Fax: 617 252 1902  
Email: thudson@genome.wi.mit.edu

Primer A: AGGCGACAAAGTCACAGCT  
Primer B: TCTCCGAAATTCATGTGG  
STS size: 132  
PCR Profile:  
Presoak:  
Denaturation:  
Annealing: 56 degrees C  
Polymerization:  
PCR Cycles: 35  
Thermal Cycler:  
Protocol:  
Template: 10 ng  
Primer: each 5 pm  
dNTPs: each 4 mM  
Tag Polymerase: 0.025 units/ul  
Total Vol: 20 ul

Buffer:  
MgCl2: 1.5 mM  
KCl: 50 mM  
Tris-HCl: 10 mM  
pH: 9.3

FEATURES  
Derived from dbEST (genbank accession 765854).  
Location/Qualifiers  
source  
1..533  
/organism="Homo sapiens"  
/map="461.5 CR from top of Chrl linkage group"  
STS  
primer\_bind 34..165  
primer\_bind complement(145..165)  
BASE COUNT 140 a 123 c 136 g 128 t 6 others  
ORIGIN

Query Match 4.5%; Score 73; DB 200; Length 533;  
Best Local Similarity 70.1%; Pred. No. 5.24e-89;  
Matches 131; Conservative 0; Mismatches 55; Indels 1; Gaps 1;

Db 250 tggacaagaggttaagcctggttcagagattccagtgatcatcattgtaata 309  
|||||  
CP 1137 tggatgagggggtatggcccttggttgcagagaaacccgcaaaatctaaatcact 1078  
|||||  
Db 310 gaccagatcatgtctctccaggtttaactttaagaactgaacctgtgtctccata 369  
|||||

CP 1077 GCCCAGACCATGAGCCGCCAGTCCTTCTGCTTCAGATAGTGCACCTGTTTGAAG 1018

Db 370 ctctcacatcatatagccaccactgttcccttgactcgtgagggagacctg 429  
|||||

CP 1017 CTCTCAGATCATCAAGCCACCACCT-GGTTGTCCGGAAGATGAGGACCTCTGTG 959

Db 430 atccctg 436  
|||||

CP 958 ATCTGTG 952

RESULT 10  
LOCUS M40107 311 bp mRNA EST 22-JAN-1996  
DEFINITION yw72b09.t1 Homo sapiens cDNA clone 257753 5' similar to gb:M80927  
CARTILAGE GLYCOPROTEIN-39 PRECURSOR (HUMAN).  
ACCESSION M40107  
NID g1163652  
KEYWORDS EST.  
SOURCE human clone-257753 primer-T7 library=Soares placenta 8to9weeks  
2NBHP8to9M vector=PT73D (Pharmacia) with a modified polylinker  
host-DH10B (ampicillin resistant) Raitel-Not I Raitel-Eco RI two  
placentae: one from 8 weeks and another from 9 weeks post  
conception. 1st strand cDNA was primed with a Not I - oligo(dt)  
primer [5'-TGTTACCAATCTGAGTGGAGGCGCCGATTTTCTTTTCTT-3'],  
double-stranded cDNA was size selected, ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I and Eco  
RI sites of a modified pT73 vector (Pharmacia). Library  
constructed by Bento Soares and M. Fatima Bonaldo.

ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Eumetazoa; Bilateria; Coelomata;  
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;  
Sarcopterygii; Choonata; Tetrapoda; Amniota; Mammalia; Theria;  
Eutheria; Archonta; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 311)  
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,  
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,  
Treviski, E., Waterston, R., Williamson, A., Wohlmann, P. and  
Wilson, R.  
The WashU-Merck EST Project  
Unpublished (1995)

TITLE  
JOURNAL  
COMMENT

Contact: Wilson RK  
WashU-Merck EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu  
High quality sequence stops: 268  
Source: IMAGE Consortium, LNL  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
Location/Qualifiers  
source  
1..311  
/organism="Homo sapiens"  
/clone="257753"  
<1..>311  
mRNA  
BASE COUNT 93 a 81 c 68 g 68 t 1 others  
ORIGIN

Query Match 4.3%; Score 71; DB 79; Length 311;  
Best Local Similarity 65.8%; Pred. No. 3.96e-85;  
Matches 156; Conservative 0; Mismatches 80; Indels 1; Gaps 1;

Db 69 ctggcaggtgtagtgcctgctctccaggaagatcgtctacaactggttg 128  
|||||  
QY 19 CTGGCAGGTTTCATGCGCTCTCTATGATCCATGGGCGCTCTGCAAACTGCTG 78  
|||||

Db 129 ctacttccaactgtgtccagcagcaggaacaggaanaa-ttaccctcagaatat 187  
|||||

QY 79 CTACTTCAACCACTGGCCAGTAGACAGAGGGAGGCTGCTTCGTCGCAAGACTT 138  
|||||



RESULT	13	R55530	396 bp	mRNA	EST	22-MAY-1995
LOCUS		yj79c12.r1	Homo sapiens cDNA clone 154966 5'		similar to gb:M80927	
DEFINITION		CARTILAGE GLYCOPROTEIN-39 PRECURSOR (HUMAN); .				
ACCESSION		R55530				
NID		9824825				
KEYWORDS		EST.				
SOURCE		human clone=154966 library=Soares breast 2mHBast vector=PT7n3D (Pharmacia) with a modified polylinker host=DH10B (ampicillin resistant) primer=M13Rpi Rsite1-Not I Rsite2-Eco RI Adult female. 1st strand cDNA was primed with a Not I - 0190(dt) primer [5' TGTTACCAGTCACTGAAAGGGAGCGGCCCTTTTATTTTTTTT 3'] , double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of modified pTZ73 vector (Pharmacia). Library went through one round of normalization to a Cot = 230. Library constructed by Bento Soares and M.Fatima Bonaldo.				
ORGANISM		Homo sapiens Eukaryota; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 396) Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kueba,T., Le,M., Lennon,G., Marra,M., Patson,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trivaskis,E., Waterston,R., Williamson,A., Woldmann,P. and Wilson,R.				
REFERENCE		The Mashu-Merck EST Project Unpublished (1995)				
AUTHORS		Contact: Wilson RK Mashu-Merck EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu High quality sequence stops: 284 Source: IMAGE Consortium, LLNL This clone is available royalty-free through LLNL; contact the IMAGE Consortium ( <a href="mailto:info@image.llnl.gov">info@image.llnl.gov</a> ) for further information.				
TITLE	JOURNAL	Location/Qualifiers 1..396 /organism="Homo sapiens" /clone="154966"				
COMMENT		BASE COUNT 88 a 106 c 109 g 87 t 6 others ORIGIN				
FEATURES	Source	Query Match 4.0%; Score 65; DB 134; Length 396; Best local Similarity 68.3%; Pred. No. 1,28e+73; Matches 114; Conservative 0; Mismatches 53; Indels 0; Gaps 0;				
Db	125	caggcttgatgtcttcgtgcgtccacagttgctctgcatacaaacagtgcctgaact 184 				
Oy	24	caggtttcatgttctctcgatgatgattccatgggctcttcgcaaaatggcttgcctact 83 				
Db	185	acacagcagtcgtccacagtaaccgganngaatgtagtgagctgttcccagatgcccttgacc 244 				
Oy	84	TCACCAACTGGGCCCAGTGACAGACAGGGGAGGCGTCGCTTCGCCAAGAGACTTGAGACC 143 				
Db	245	gnntctctgtatccacatcatcatcaagcttgccaatatagaacaac 291 				
Oy	144	CCAGCCTTTGCACCCACCTCATCTAGCGCTTGCGTGCGCATGTGACCAAC 190 				
LOCUS	14	M10705	306 bp	mRNA	EST	26-APR-1996
DEFINITION		ma47g07.r1 Soares mouse p3BMF19.5 Mus musculus cDNA 5'			similar to	
ACCESSION		gb:M80927			CARTILAGE GLYCOPROTEIN-39 PRECURSOR (HUMAN); .	
		M10705				

NID	KEYWORDS	EST.
91285179	house mouse.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	
	Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae; Murine; Mus.	
REFERENCE	1 (bases 1 to 306)	
AUTHORS	Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, D., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.	
TITLE	The WashU-HMI Mouse EST Project	
JOURNAL	Unpublished (1996)	
COMMENT	Contact: Marra M/Mouse EST Project WashU-HMI Mouse EST Project Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watsn.wustl.edu This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Seq primer: mob.REGA+ER High quality sequence stop: 172.	
FEATURES	NCBI g1: 1285179	
SOURCE	Location/Qualifiers 1..306 /organism="Mus musculus" /note="Vector: pT730 (pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - Oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGCATTTTATTTTATTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Failla Bonaldo." /clone_lib="Soares mouse p3NMF19.5" /dev_stage="19 weeks" /lab_host="DH10B (ampicillin resistant)" 1..>306	
BASE COUNT	78 a 76 c 85 g 67 t	
ORIGIN	mrna	
Query Match	3.7%; Score 60; DB 215; Length 306;	
Best Local Similarity	72.1%; Pred. No. 3.39e-64;	
Matches	98; Conservative 0; Mismatches 38; Indels 0; Gaps 0;	
Db	78 agaagcttcctcgtcctaccagaagcaccacggtggtggtatagagcaagagagtg 137	
960	agaagcttcctcgtcctaccagaagcaccacggtggtggtatagagcaagagagtg 1019	
Db	138 tcaaaacaaggttggtctcgaagaggaagaacgtgcagagcatcggtgtggcac 197	
1020	tcaaaacaaggttggtctcgaagaggaagaacgtgcagagcatcggtgtggcac 1079	
Db	198 tggatttgatgatctt 213	
1080	tggacttagatgactt 1095	
RESULT	15	
LOCUS	W10705 306 bp mrna EST 05-SEP-1996	
DEFINITION	ma7g907.f1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 313884 5'	
ACCESSION	W10705	
NID	91285179	
KEYWORDS	EST.	

SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Mus.  
REFERENCE 1 (bases 1 to 306)  
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,  
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,  
Schellenberg,K., Stepcoe,M., Tan,F., Underwood,K., Moore,B.,  
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and  
Waterston,R.  
TITLE The WashU-HMI Mouse EST Project  
JOURNAL Unpublished (1996)  
COMMENT  
Contact: Marra M/Mouse EST Project  
WashU-HMI Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LIND ; contact the  
IMAGE Consortium (info@image.lind.gov) for further information.  
Seq primer: mob REGA+ET  
High quality sequence stop: 172.  
Location/Qualifiers  
1..306  
/organism="Mus musculus"  
/note="Vector: pT73D (Pharmacia) with a modified  
polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
was primed with a Not I - oligo(dT) primer [5'  
TCCTACCAATCTGAAGTGGAGCGCGCGCATTTTCTTTTCTTTT 3'] ,  
double-stranded cDNA was size selected, ligated to Eco RI  
adapters (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of a modified pT73 vector  
(Pharmacia). Library went through one round of  
normalization to a Cot = 5. Library constructed by Bento  
Soares and M.Fatima Bonaldo. RNA was kindly provided by  
Dr. Minoru Ko (Wayne State University)."  
/clone\_id="313884"  
/dev\_stage="19.5 dpc total fetus"  
/lab\_host="DH10B (ampicillin resistant)"  
<1..>306  
BASE COUNT 78 a 76 c 85 g 67 t  
ORIGIN  
mRNA  
BASE COUNT 78 a 76 c 85 g 67 t  
ORIGIN  
Query Match 3.7% Score 60; DB 174; Length 306;  
Best Local Similarity 72.1% Pred. No. 3.39e-64;  
Matches 98; Conservative 0; Mismatches 38; Indels 0; Gaps 0;  
Db 78 agaaggttccttcgcatacaagggcaacagctgggtgggtatgaagacagaagagtg 137  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 960 AGAAGTGCCTTCATCTTCGCGACACCAAGTGGGTGGCTTGATGATGTGAGAGCT 1019  
Db 138 tcaaaaacaggttggttcttcgaagagaagaagctggcagagacatcggtgtggcac 197  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 1020 TCAAAACCAAGCTACTATCTGAACAGCAAGGACTGCGGGCCATGCTGTGGGCAC 1079  
Db 198 tggatttgatgattt 213  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 1080 TGGACTTAGATGACTT 1095

Search completed: Sun Jun 29 20:18:52 1997  
Job time : 1181 secs.



\*\*\*\*\*  
 Release 2.1D John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
 Distribution rights by IntelliGenetics, Inc.  
 \*\*\*\*\*

(TM)

MPsrch\_n n.a. - n.a. database search, using Smith-Waterman algorithm  
 Run on: Sun Jun 29 19:29:43 1997; Maspar time 1518.90 Seconds  
 Tabular output not generated. 1197.646 Million cell updates/sec

Title: >US-08-663-618A-1  
 Description: (1-1636) from US08663618A.seq  
 Perfect Score: 1636  
 N.A. Sequence: 1 CATGCTGCGGCTCTGTGGCT.....CCCTCTCCCAAAAAA 1636  
 Comp: GTACGACGCGACACACCGGA.....GGGAGAGGCTTTT

Scoring table: TABLE default  
 Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 333249 segs, 555961234 bases x 2

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database:

emb1-new11  
 1: BCT 2: FUN 3: GEN 4: HUM1 5: HUM2 6: HUM3 7: INV1 8: INV2  
 9: INV3 10: INV4 11: INV5 12: INV6 13: INV7 14: ORG 15: MAM  
 16: VRT 17: PLN 18: PRO1 19: PRO2 20: ROD 21: SYN 22: UNC  
 23: VIR1 24: VIR2

Database:

genbank97  
 25: BCT1 26: BCT2 27: BCT3 28: BCT4 29: BCT5 30: BCT6 31: BCT7  
 32: BCT8 33: BCT9 34: GEN1 35: GEN2 36: HGT 37: INV1 38: INV2  
 39: INV3 40: INV4 41: INV5 42: INV6 43: INV7 44: INV8 45: INV9  
 46: MAM1 47: MAM2 48: MAM3 49: VRT1 50: VRT2 51: VRT3 52: PAT1  
 53: PAT2 54: PAT3 55: PAT4 56: PHG 57: PLN1 58: PLN2 59: PLN3  
 60: PLN4 61: PLN5 62: PLN6 63: PLN7 64: PLN8 65: PLN9 66: PLN10  
 67: PRI1 68: PRI2 69: PRI3 70: PRI4 71: PRI5 72: PRI6 73: PRI7  
 74: PRI8 75: PRI9 76: PRI10 77: PRI11 78: PRI12 79: PRI13  
 80: PRI14 81: ROD1 82: ROD2 83: ROD3 84: ROD4 85: ROD5 86: ROD6  
 87: ROD7 88: ROD8 89: STR 90: SYN 91: UNA 92: VRL1 93: VRL2  
 94: VRL3 95: VRL4 96: VRL5 97: VRL6 98: VRL7 99: VRL8 100: VRL9

Database:

genbank-new11  
 101: BCT 102: GEN 103: INV1 104: INV2 105: MAM 106: VRT  
 107: HGT 108: PLN 109: PRI1 110: PRI2 111: ROD 112: SYN  
 113: UNA 114: VRL  
 u-emb148\_97  
 115: part1 116: part2

Database:

Statistics: Mean 11.639; Variance 4.516; scale 2.577

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

# SUMMARIES

Result Query Match Length DB ID Description Pred. No.

1	1603	98.0	1633	74	HSU29615	Human chitotriosidase	0.00e+00
2	184	11.2	1741	76	HUMHA3G	Human glycoprotein MR	3.37e-154
3	159	9.7	1418	74	HU549835	Human YKL-39 precursor	5.02e-128
4	159	9.7	1434	74	HSU58514	Human chitinase precu	6.03e-128
5	157	9.6	1500	74	HSU58515	Human chitinase (HMT)	6.03e-126
6	156	9.5	1733	48	SSGP38KD	S. scrofa 38kDa hepari	6.59e-125
7	156	9.5	1733	48	SSU19900	S. scrofa 38kDa hepari	6.59e-125
8	156	9.5	1733	15	SSGP38KD	S. scrofa 38kDa hepari	6.59e-125
9	135	8.3	1616	81	MMBRP39	M. musculus mRNA for B	3.23e-103
10	123	7.5	1506	20	MMD757	House mouse: Musculus	5.92e-91
11	123	7.5	1506	20	D87757	House mouse: Musculus	5.92e-91
12	105	6.4	2022	48	SSU43490	House mouse: Musculus	8.60e-73
13	97	5.9	2034	47	OAUI6719	Sus scrofa oviductal	7.86e-65
14	93	5.7	1596	47	OAUI7988	Ovis aries oviductal	6.99e-61
15	93	5.7	2228	67	BADEDOG	Papio hamadryas andubi	6.99e-61
16	93	5.7	2228	15	PAEDDOG	Papio hamadryas andubi	6.99e-61
17	93	5.7	2525	20	MMOSC	Mouse mRNA for oviduc	6.99e-61
18	93	5.7	2525	85	MOSOGP	Mouse mRNA for oviduc	6.99e-61
19	89	5.4	2198	73	HSU09550	Human oviductal glyco	5.90e-57
20	85	5.2	1994	46	BOV95BOGP	Bovine mRNA for 95 kD	4.70e-53
21	69	4.2	2387	20	MAOSG	Hamster mRNA for ovid	9.34e-38
22	69	4.2	2387	81	HAMOGP	Hamster mRNA for ovid	9.34e-38
23	59	3.6	2452	41	MAU15048	Mesocricetus auratus	1.61e-28
24	52	3.2	2452	41	MSU02270	Manduca sexta chitina	2.91e-22
25	45	2.8	443	72	HSOGP03	Mus musculus chitinas	3.14e-16
26	45	2.8	994	83	MMU56900	Human oviductal glyco	9.92e-14
27	42	2.6	3060	29	KURCHISHI	Kurthia zopfi gene f	6.55e-13
28	41	2.5	506	72	HSOGP07	Human oviductal glyco	2.71e-11
29	39	2.4	2078	44	PEURCHT1	Petaurus japonicus adu	2.71e-11
30	40	2.4	2401	25	BACCHIA3	B. circulans chitinase	4.25e-12
31	36	2.2	25083	7	CEC04F6	Caenorhabditis elegans	6.27e-09
32	36	2.2	25083	38	CEC04F6	Caenorhabditis elegans	6.27e-09
33	34	2.1	1322	7	BPUS9689	Brugia pahangi microf	2.12e-07
34	34	2.1	1322	37	BPUS9689	Brugia pahangi microf	2.12e-07
35	34	2.1	1322	103	BPUS9689	Brugia pahangi microf	2.12e-07
36	34	2.1	1654	37	BRPCHIT	Brugia pahangi microf	2.12e-07
37	33	2.0	490	44	OVU14639	Onchocerca volvulus c	1.19e-06
38	32	2.0	1364	7	BPUS9690	Brugia pahangi microf	6.54e-06
39	32	2.0	1364	103	BPUS9690	Brugia pahangi microf	6.54e-06
40	32	2.0	1364	37	BPUS9690	Brugia pahangi microf	6.54e-06
41	32	2.0	1700	32	SMCHTB	Serratia marcescens c	1.19e-06
42	33	2.0	1836	32	SHCHTB	S. marcescens (BUL200)	1.19e-06
43	31	1.9	1320	54	II4183	Sequence 13 from pate	3.48e-05
44	31	1.9	1364	54	II4182	Sequence 13 from pate	3.48e-05
45	31	1.9	1405	54	II4180	Sequence 9 from paten	3.48e-05

## ALIGNMENTS

RESULT 1  
 LOCUS HSU29615 1633 bp mRNA PRI 04-NOV-1995  
 DEFINITION Human chitotriosidase precursor mRNA, complete cds.  
 ACCESSION U29615  
 NID 91050957

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (21-JUN-1995) Rolf G. Boot, Department of Biochemistry,  
 AMC, E.C. Slater, Institute of Amsterdam, Weibergdreef  
 15, 1105 AZ, Amsterdam, The Netherlands



FEATURES  
 source  
 Location/Qualifiers  
 1..1633  
 /organism="Homo sapiens"  
 /clone\_lib="Ch1.1"  
 /cell\_type="macrophage"  
 /cell\_type="macrophage"  
 13..75  
 13..1413  
 /codon\_start=1  
 /product="chitotriosidase precursor"  
 /db\_xref="pid:g1050958"  
 /translation="MVRVAMAGFVLLMIIPWSAKLYCYFTTMAQYRNGARELPK  
 DLPISCTHLIYAFAGMTNHOISTEMNDELIOENGLKKNPKILKILAIAGWNFG  
 TOKFTDMVATANNROTIVNSAIRFLAKYSFEDGLDMWEPGSGSPADYKERFTLVQ  
 DLANFOEAOISGKRELLSAIPAGOTYVDAGYEVDAKIAONLDPVNTMAYDFHGSW  
 EKVTHNSPLRYKROESGASALINDAAYOOLKOTRPAKILHPTYGRSFTLASS  
 SDTRGAPATGSGTREPPTKEGMLAIYEVSGMKATKRIODQKPTIFRNOWVGF  
 DVESEKIVSYLKLKGLGAMVWLIDLDFAGFSCNDRPLIOTLRQELSLPLPS  
 GTPLEVEPRGPSPSEPHCPSPGODTFCKGADGLYPNRRERSFYSCAAGRLFOOSC  
 PGLVFSNCKCKCTWN"  
 75..1410  
 /note="human chitinase"  
 /product="chitotriosidase"  
 1605..1610  
 1633  
 /note="10 A nucleotides"  
 BASE COUNT 354 a 489 c 443 g 347 t  
 ORIGIN  
 Query Match 98.08; Score 1603; DB 74; Length 1633;  
 Best Local Similarity 99.88; Pred. No. 0.00e+00;  
 Matches 1621; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

Db 612 cgcacgaacctgatttctacacctatagcctcactccatcctcctcttggaaga 671  
 QY 601 ccccccagAACCTGGATTGTGCAACCTTATGGCCCTACAGCTTCATGGCTCTGGGAGAA 660  
 Db 672 gtccagggacataaacaagcccccctacagaagaggaagaagaagatgctgacgacga 731  
 QY 661 ggtcACGGGACATTAACAGCCCCCTCTACAAAGAGCAAGAGAGAGTGGTGCAGCAGCAG 720  
 Db 732 cctcaacgtggtgtcgtcgtctgacacagtggtgtgcaagaagggagaccctgcagaagct 791  
 QY 721 CCTCAACGTGGATGCTGCTGTGCAACAGTGGCTGGAAGAGGGGACCCCTGCGCAGAGCT 780  
 Db 792 gatcctctgcaagcctacactacagagcgcctcctcaactgagcctcctcaagacacga 851  
 QY 781 GATCCTTGGAATGCTACTACTAGGAGCGCTCTTACACTGGGCTCTCTATCAGAACACAG 840  
 Db 852 agtgggggccccagccacagaggtctgcactcacaagccccctacacaaaggaagaaggat 911  
 QY 841 AGTGGGGGCCACAGCCACAGGGGTGCGACTCCAGGCCCTTACCAAGGAAGAGGAGAT 900  
 Db 912 gcttgacctacttgaaagtctgctcctctggaagggggccacaaacagagaatccagatca 971  
 QY 901 GCTGGCTTACTATGAAGTGTGCTCTGGAAGGGGGCCACAAACAGAGATCCAGATCA 960  
 Db 972 gaaggtgcccctacatctctccgggaacaacagtggtggtggtcttgatgtatgtagagact 1031  
 QY 961 GAAGGTGCCCTTACATCTTCCGGGACACACAGTGGGTGGCTTGTATGATGTGAGAGACTT 1020  
 Db 1032 caaaacaaagtcagatctcgaagcaagaaggaagctggcgaggaatgcttgagcact 1091  
 QY 1021 CAAACCAAGGTGAGCTATCTGAGCAGAAAGGAGCTGGGGCCGAGTGTGGGCACT 1080  
 Db 1092 ggaactgaatgacttggcgggtcttcctcgcgaacgaagggcgatccccctatacgaac 1151  
 QY 1081 GGACTTGAATGACTTGGCCGCTTCTCTGCAACCAAGGGCCGATACCCCTCATCCAGAC 1140  
 Db 1152 gctacggcaggaacttgactctccatctacttgctctcgaagcaccacgaagcttgaaagctc 1211  
 QY 1141 GCTAGCGGAGAACTAGTCTTCATCTTGGCTTTCAGGACCCAGAGCTTGAAGTTCC 1200  
 Db 1212 aaaaacaggtcagccctctgaacctgaagcagtgagccccagcccttggaacagacgtctctg 1271  
 QY 1201 AAAACAGGTGACCCCTCTGAACTGAGCATGGCCCAACCCCTGGAACAAGACACGTTCTG 1260  
 Db 1272 ccaagggcaagcttgatgggtctctatcccaactcctctgggaaggttcaagcttcaagctg 1331  
 QY 1261 CCAAGGCAAGCTGATGGGCTCTATCCCAATCCTGGGAACGGTCCAGCTTCTTACAGCTG 1320  
 Db 1332 tgcagcgaggcggtctgttcacgaacaagctgcccgaacagcgcctggtgttcagcaactctg 1391  
 QY 1321 TGCAGCGGGGGCGGCTGTTCACGCAAAAGCTGCCGACAGGCTCGTGTTCAGCAACTCTCG 1380  
 Db 1392 caaatgctgacacctggaattgagtcg-aaagccccctccagtcg-agcttgaggctggg 1449  
 QY 1381 CAATGCTGACCTGGAATTTGAGTGCCTTAAAGCCCTCCAGTCCAGCTTTGAGGCTGGG 1440  
 Db 1450 cccaagatctactctaaagctgctcctctggtttcc-tgggggcgcgaatttgctcct 1508  
 QY 1441 CCCAGGATACCTTACAGCTGCTGCTCTGCTGGTTTCCCTGGGGGCGCGCAATGTGCTCTCT 1500  
 Db 1509 gcaagccttctggtgtctcctcttaccagagcttctcctcctcaagccttgctcctcct 1568  
 QY 1501 GCAAGGCTTTCGTGTGTCTTCTCTTATCCAGGCTTTCCTCCTCAGACCTTGCCTTCCTT 1560  
 Db 1569 ttcttggtctccttggtgctgcccccttcaacttgcaaaataaactcttggttctgccccct 1628  
 QY 1561 TTCTCTGGGTCTCCTGGGCTGCCCCCTTTCACCTTGCAAAATAATCTTGTGTGTGCCCT 1620  
 Db 1629 cttc 1632  
 QY 1621 CTTTC 1624

```

RESULT 2
LOCUS HUMHA3G 1741 bp mRNA PRI 24-JAN-1994
DEFINITION Human glycoprotein mRNA, complete cds.
ACCESSION M80927
NID 9348911
KEYWORDS glycoprotein.
SOURCE Homo sapiens
ORGANISM Homo sapiens cDNA to mRNA.
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 1741)
Hakala,B.E., White,C. and Recklies,A.D.
TITLE Human cartilage gp-39, a major secretory product of articular
chondrocytes and synovial cells, is a mammalian member of a
chitinase protein family
JOURNAL J. Biol. Chem. 268 (34), 25803-25810 (1993)
MEDLINE 94064658
FEATURES
source location/Qualifiers
1..1741
/organism="Homo sapiens"
/cell_type="chondrocyte"
/tissue_type="cartilage"
1..71
5'UTR 72..134
sig_peptide /codon_start=1
72..1223
CDS /note="articular 39kDa glycoprotein"
/codon_start=1
/evidence="experimental"
/product="glycoprotein"
/db_xref="pid:9348912"
/translation="MGVKASQGFVYLVLQCSAYKLYCYTSQYREDSCEPD
ALDRLCRTHIYSEFANISNDHIDTENDVLTGKMLTKRNPMLKTLISYGVNNG
SQRPSKIASNTQSRFTIKSVPELRTGFGDLAMLYPERDQHPFTLIKEMKAE
FIKENAPGKQDLISALASAGKVIDSSYDIASQHLDFISIMYDHWGMRGTGH
HSPLEFGQEDASPDREFTNTDYGMLRLGAPASLVAGIPFGSFTLASSENGVA
PISGPIRGFTKREAGTLAYEICDFELAGAPASLVAGIPFGSFTLASSENGVA
KSKVYIKDRQLAGMVMALDLDPDQSGFCODLRFPLTKAKDLAT"
mat_peptide 135..1220
/codon_start=1
/product="glycoprotein"
3'UTR 1224..1741
polyA_signal 1715..1720
polyA_site 1741
BASE COUNT 416 a 510 c 447 g 368 t
ORIGIN
Query Match 11.2% Score 184 DB 76: Length 1741:
Best Local Similarity 63.8% Pred. No. 3.57e-154:
Matches 701: Conservative 0: Mismatches 373: Indels 24: Gaps 10:

```

```

QY 324 ATATGTATACCAAGGCGCAACACGTCAGACTTTGTCAACTCGGGCATCAGGTTTCGC 383
DB 454 gcaaccatgcttcatggtctgagacctgctgtcttaacctgga---c---ggag-- 505
QY 384 GCMAATACAGCTTTGACGGCTTGACCTTGAGTGGAGTACCAAGAACCCAGGGAGGCC 443
DB 506 -----agacaaagacatttacacacctatcaagaaatgaagccgaattataa 558
QY 444 CTGCGGTACACAGAGACCCCTTACACACCTTGTCACAGACTTTGGCAATGCTTCCACG 503
DB 559 aggaagccag--c--cagggaaaaagcagctctctcagcgcagcagctgtctgaggga 615
QY 504 AGGAAGCCAGACCTCAGAGGAAGAAACGCTTCTGTGATGACACAGGTTCCAGTGGCC 563
DB 616 aggtacacattgacagcagctatgacattgccaatatccacaacctgattcaata 675
QY 564 AGACCTATGTGATCTGTGATGAGAGTGAGCAAAATCCCGAAGACCTGGATTGTTGTA 623
DB 676 gcatcatgacctaagatttcatgtgagcctggtggtggaacacagggccatcagtcgcc 735
QY 624 ACCTTATGCTTACAGACTTCCATGCTCTTGGAGAGAGTTCAGGAGACTTAACAGCCCC 683
DB 736 tgtccgaagtcagagagatgcaagtccttgacagattcagacacatgactatgtgtg 795
QY 684 TCTACAGAGCGCAAGAGAGAGTGTGACAGCAGCAGCCTCAAGTGAATGCTGCTGTC 743
DB 796 gttacatgttgaagctggggctcctgccaagtaagctgtgtatggtgcatcccaacctg 855
QY 744 AACAGTGTGCTGCAAGAGGAGCCCTGCAGCAAGCTGATCCTTGATGCTTACTACG 803
DB 856 ggaagagctctctctgtgtcttc-tgagactg--ggttggaaccccaactcagac 912
QY 804 GACGCTCTTCACTGACCTGCTCTCTCATCAGACCAAGAGTGGGGCCCGACCAAGGGT 863
DB 913 cgggaattcagagccggttcaacaagagagcagggaccttccactactatgagatctgt 972
QY 864 CTGGCACTCCAGGCCCTTCAACCAAGAGAGGAGATGCTGGCTACATGATGAGTCTGT 923
DB 973 acttctccgcggagccacagatcatalgaacctgcgcagcaggttccctatgccaaca 1032
QY 924 CTTGGAAG-G-GG-GCCACCAACAGAGATCCAGAGATCAGAAAGTGCCTACATCTTC 980
DB 1033 agggcaaccagtggttagatcagcagcaggaagcgtcaaaagcgaagtgacgtacc 1092
QY 981 GGGACACACAGTGGGTGGCTTGTGATGTGAGAGAGCTTCAAAACCAAGACAGCTATC 1040
DB 1093 tgaagatagcagctgagcagcagcagctgataggccctgagcctgagctgagctccag 1152
QY 1041 TGAAGCAGAGGAGACTGGGCGGGCCATGCTGTGGGCACTGAGATTAGATGACTTGGCC 1100
DB 1153 gctcctctgcggccag 1170
QY 1101 GCTTCTCTGCAACACAGG 1118

```

RESULT 3  
 LOCUS HSU49835 1418 bp mRNA PRI 25-JUL-1996  
 DEFINITION Human YKL-39 precursor mRNA, complete cds.  
 ACCESSION U49835  
 NID 91457940  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;  
 Vertebrata; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE  
 AUTHORS Hu,B., Trinh,K., Figueroa,W.F. and Price,P.A.  
 TITLE Isolation and sequence of a novel human chondrocyte protein related  
 to mammalian members of the chitinase protein family  
 JOURNAL J. Biol. Chem. 271 (32), 19415-19420 (1996)  
 MEDLINE 96325055  
 REFERENCE  
 PRICE,P.A.

TITLE Direct Submission  
JOURNAL Submitted (23-FEB-1996) Paul A. Price, Biology, 0322, University of California at San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0322, USA

FEATURES  
source Location/Qualifiers  
1..1418  
/organism="Homo sapiens"  
/tissue\_type="articular cartilage"  
36..1193  
/note="Similar to human glycoprotein encoded by Genbank Accession Number M80927, human chondroitinase precursor encoded by Genbank Accession Number U29615, human ovoductal glycoprotein encoded by Genbank Accession Number U09550, and to mouse secretory protein YM-1 precursor encoded by Genbank Accession Number S27879"  
/codon\_start=1  
/product="YKL-39 precursor"  
/db\_xref="PID:91457941"  
/translation="MDQSLMAGVYVLLLLQGSAYKLVCFYTNMSODROPEKFTPE NIDPFLCSHLIYSFASILENNKVIIDKSEYMLYQYINSLTKNPKLILISGYLFG SKGFHPWVDSSTSLRLEINSLILELRNHDGDLVSWIYDOKENTHEVLIHLAIA FQDETSTKRELITAGSAGROMDINSYOVERLAIDLFINLSDFHGSMEKPII TGHNSPLSKGMORBPSSSYNVEAYVYHKGKPSKRYMGPIYTHSFTLSAETT VGAPASGPAAGPTTESGFLATIEICQFLKAKITLDOQVYAVKGMQWGYDDV KSMETKYQFLKNNLNGAMIMSIDMDFTGKSCNQPYPYVQAVKRLSLGL"

sig\_peptide 99..1190  
mat\_peptide 36..98  
polyA\_site /product="YKL-39"  
BASE COUNT 370 a 357 c 337 g 354 t  
ORIGIN

Query Match 9.7%: Score 159; DB 74; Length 1418;  
Best Local Similarity 63.0%; Pred. No. 5,026-118;  
Matches 711; Conservative 0; Mismatches 350; Indels 27; Gaps 10;

D 53 ctggagagtgtagtgctgtgctgtgcttcacaggaagatgcgtacaaactggttg 112  
Y 19 CTGGGACAGTTTCAMGGTCTGCTGATGATCCCATGGGCTCTGTCGAAAACGTGTG 78

D 113 ctacttcaacaactggtccagagcaggaagaccaggaataatccacctgagaatat 172  
Y 79 CTACTTCACCAACTGGGCGCCAGTACAGACAGGGAGGCTGCTGCTGCCCAAGACACT 138

D 173 tgaccacctatgctctctcatctatctatctatctgcccagctcggaaacaaaggt 232  
Y 139 GGACCCAGCCTTGTGCACCCACCTCATCTACGCTTCCTG6CATGACCAACCACT 198

D 233 tatcatcaagacaagaatgaagtgatgctctacacagacatcaacagctcaaaacaa 292  
Y 199 GAGCCACGCTGATGATGAGTACAGACACCTCTACCAAGAGTTCAATGGCTCGAAGAAAT 258

D 293 gaatcccaactgaataatctctctgtgcatatggagaggtactgttggttccaaaggtc 352  
Y 259 GAATCCCAAGCTGAAGACCTCTGTAGCATGGAGGCTGGAATTCGGCACCTCAAGATTT 318

D 353 caaccctatggtgattcttctacatccagcttggaattcaattcaatccaatccgtt 412  
Y 319 CACAGATATGTGATGACGACGACCAACACCGTTCGTCACCTGGCCATAGAGTT 378

D 413 tctagagacaataacttgatgactgagtgaactggaatctccagc--atc-aggaa 469  
Y 379 TCTGCGCAAAATACAGCTTTGACGGCTTGACCTTACTGGAGTACCCACAGAGCAAGG 438

D 470 -agaa-----a-acatcat--ttcactgtgctgattcatgattagagaagacct 517  
Y 439 GAGCCCTCCGTAGACAGAGAGCGCTTCACCAACCTGTGACAGACTTGCCCAATGCTT 498

D 518 tcagaagagacttcacaaataccacaaagaaaggtctctctgactggggcgtctcgc 577  
Y 499 CCAGAGGAAGCCCAAGACCTCAGGAGGAAGAACGCTCTTCTTGTAGTGCAGGCGTTCACG 558

D 578 aggaagacaatgatgtatatacagctatacaagtgaagaactcgcaaaagatcgtgatt 637  
Y 559 TGGGCAACACCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 618

D 638 catcaactctcgtctcttcttacttcaatcattggtcttgggaaagcccttactatgcca 697  
Y 619 TGTCAACCTTATGACCTACGACCTTCCATGCTCTTGGGAAAGG-----TCACGGGACA 672

D 698 caaagccctctggaagaggggtggtcgaggaagggccagtcctactacaaatgga 757  
Y 673 TAAAGCCCTCTTACAGAGAGGACAGAGAGAGTGGTGCAGCAGCCTCAACGTGGA 732

D 758 atatgctgtggtgactgatacataaaggaatgccatcagaagaagtggtgactatgcat 817  
Y 733 TGTGCTGTGCAAGAGTGGTGCAGAGAGGAGGAGCCCTCCAGCAGAGATGATCTGGCAT 792

D 818 ccccaataatggatcctccttcaacatggtcctc--cagaacacacgtggggccc 874  
Y 793 GCTTACCTACGAGCGCTTTCACACTGCGCTCCTCATCAGACACAGAGTGGGGCCCC 852

D 875 tgcctcggccctgagctgctgagccatcaacagatcttcaggtcctcgtgactata 934  
Y 853 AGCCACAGGCTGTGCACTCAGCCCTTCCACCAAGAGAGAGAGTGTGCTGCTACTA 912

D 935 tgaatctgcccagttccttgaagaagagccaagatcacgcgcctccagatcacaggttcc 994  
Y 913 TGAAGTCTGC--TCCAGAGAGGGGGCCACCAACAGAGATCAGAGATCAGAGTGC 969

D 995 ctacgcagtcgaaggggaacagtggtggtgctatgatatgatatgatatgatatgatat 1054  
Y 970 CTACATCTTCGCGGACACACAGTGGTGGTGTGATGATGATGATGATGATGATGATGAT 1029

D 1055 ggttcagttcttaagaagaattaaacctggagagagagagagagagagagagagagag 1114  
Y 1030 GGTGAGCTATGTGAAG 1089

D 1115 tgaattcactggaatcctcctgcaacacagggcccttaccctctgtcca 1162  
Y 1090 TGACTTTGGCGGCTTTCCTGTCACAGAGGGGCGATACCCCTCATGCA 1137

RESULT 4  
LOCUS HSU58514 1434 bp mRNA PRI 24-JUL-1996  
DEFINITION Human chitinase precursor (HUMCHIT) mRNA, exon 1a form, complete cds.  
ACCESSION U58514  
NID 91439565  
KEYWORDS human.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1434)  
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.  
The WashU-Merck EST Project  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1434)  
Grossman, A., Matsuyama, T., Baker, E., Waterhouse, P., Sutherland, G. R. and Mak, T. W.  
Cloning of a novel lymphoid restricted human chitinase and localization to Ip13.3  
JOURNAL Unpublished  
REFERENCE 3 (bases 1 to 1434)  
Grossman, A., Matsuyama, T., Baker, E., Waterhouse, P., Sutherland, G. R. and Mak, T. W.  
TITLE Direct Submission  
JOURNAL Submitted (17-MAY-1996) Medical Biophysics, Ontario Cancer Institute, Princess Margaret Hospital, 610 University Avenue, Toronto, Ontario M5G 2M9, Canada

FEATURES  
SOURCE  
Location/Qualifiers  
1..1434  
/organism="Homo sapiens"  
/map="1p13.3"  
/chromosome="1"  
/tissue\_type="Soares placenta"  
/dev\_stage="8 to 9 weeks"  
/clone="EST 257753"  
1..36  
/gene="HUMTCHIT"  
1..76  
/gene="HUMTCHIT"  
/note="exon 1a; first alternatively splice exon 1; contains a kozak consensus initiator methionine"  
/number=1  
37..114  
/gene="HUMTCHIT"  
37..1209  
/gene="HUMTCHIT"  
/note="first alternatively spliced form, see GenBank Accession Number U58515 for the second alternatively spliced form"  
/codon\_start=-1  
/product="chitinase precursor"  
/db\_xref="PID:g143956"  
/translation="MGATITMDQSLMAGVYVLLLOGGSNRYCYFTNMSQDRPBG  
KTPENIDPLCSLITSFASISNNKYIINKSEVMYDITNSLTKNPKLITLLSIG  
GYLFGSKGFPMVDSSISRLFINSLILFRNHPGLDYSWIPQKNEHTPTVLIH  
ELAEAFQKDFPKSKERLLITAGVSAGROMIDNSYVEKLANKDLDFINLSFPHFSM  
EKPLITGNHSPILSKGMDRGPSSYNYVEYAVGYWIKHGMPEKVMGIPYHGSFTLA  
SAETTVGAPASGGAAGPITSSGFLAYVYICQIFGAKITRIODQOVPAAVGNQMV  
GYDVKSMERKVOFLKLNINLGAMISIDMDPTGKSCNGPIPLVQAKRSLGS"  
77..106  
/gene="HUMTCHIT"  
/number=2  
107..309  
/gene="HUMTCHIT"  
/number=3  
115..1206  
/gene="HUMTCHIT"  
/note="first alternatively spliced form, see GenBank Accession Number U58515 for the second alternatively spliced form"  
/product="chitinase"  
310..365  
/gene="HUMTCHIT"  
/number=4  
366..516  
/gene="HUMTCHIT"  
/number=5  
517..643  
/gene="HUMTCHIT"  
/number=6  
644..771  
/gene="HUMTCHIT"  
/number=7  
772..954  
/gene="HUMTCHIT"  
/number=8  
955..1071  
/gene="HUMTCHIT"  
/number=9  
1072..1433  
/gene="HUMTCHIT"  
/number=10  
1210..1433  
/gene="HUMTCHIT"  
1407..1412  
/gene="HUMTCHIT"  
1434  
/gene="HUMTCHIT"  
/note="32 A nucleotides"  
BASE COUNT 375 a 360 c 342 g 357 t

ORIGIN  
Query Match 9.7%, Score 159; DB 74; Length 1434;  
Best Local Similarity 63.0%; Pred. No. 5,02e-128;  
Matches 711; Conservative 0; Mismatches 390; Indels 27; Gaps 10;  
Db 69 ctggcagagtgatgtgtctgtctgtctccaggaggatctctcaaacctggttg 128  
19 ttt 78  
Db 129 ctacttaccacactggtccagagccgaggaaccaggaattccacccctggaatat 188  
79 ctactttaccacactggtccagagccgaggaattccacccctggaatat 138  
Db 189 tgaaccttcctatgctcctatcctatcctatcctatcctatcctatcctatc 248  
139 ggaaccttcctatgctcctatcctatcctatcctatcctatcctatcctatc 198  
Db 249 tatcatcaagagcaagagtgatgtatgtcttaccagaccatcaacagcttcaaaccaa 308  
199 gaggaccactgagtgatgacgagactcttaccagagacttcaatgacctgaagaagat 258  
Db 309 gaatcccaactgaaattctctgtccattgagaggttacctgtgttccaaagggtt 368  
259 gaatcccaactgaaattctctgtccattgagaggttacctgtgttccaaagggtt 318  
Db 369 caacctatggtgattcttctatcatcagcttggaaattcaattcaataatcctgtt 428  
319 caccagatgagtgagccagccagccagccagccagccagccagccagccagccag 378  
Db 429 tctgaggaaccatcacttctatgagctgagatgtatgaatgtatcaccag--atc-aga 485  
379 tctgaggaaccatcacttctatgagctgagatgtatgaatgtatcaccag--atc-aga 438  
Db 486 -agaa-----a-acactat--ttcactgtgctgattcatatgattgacagaagcctt 533  
439 gaggccctgctgagacagagagagagagagagagagagagagagagagagagag 498  
Db 534 tcagaaggacttcaaaaatccacaaagaagaagcttctctgactcgggcyglatctgc 593  
499 ccagcag 558  
Db 594 agggagagcaaatgattgatacagctcactcaagtgtgaaactggcaaaatctgattc 653  
559 tggcagacactatgagtgatgagctgagatgagtgagagagagagagagagagag 618  
Db 654 catcaacctctgtctcttacttccactcagagcttggagaaagcccttcaactgcca 713  
619 tgcacaccttattgagcttaccagcttccacttcttggagagagagagagagagag 672  
Db 714 caacagccctctgagcaaggggtgagcagagagagagagagagagagagagagag 773  
673 taacagccctctgagcaaggggtgagcagagagagagagagagagagagagagag 732  
Db 774 atatgctgtggtgtgagctgagagagagagagagagagagagagagagagagagag 833  
733 tgcctgctgtgag 792  
Db 834 cccacacatgag 890  
793 gccctgctgtgag 852  
Db 891 tgcctgctgtgag 950  
853 agccacag 912  
Db 951 tggatctgag 1010  
913 tgaagctgtg--tccag 969  
Db 1011 ctacgag 1070  
970 ctacgag 1029



QY	604	aaggagcttcacaaaatcccccgaagaaagcttctcttgcgcyggcggtatctcgaag	663
Db	604	aaggagcttcacaaaatcccccgaagaaagcttctcttgcgcyggcggtatctcgaag	663
QY	503	CAGGAAGCCAGACCTCAGGGAAGGAACCCCTTCTTCTTAAGTCAGACGGTTCCAGCTGGG	562
Db	664	aggcaaatgatgtgaacaacagcctacaagctatgtgaagctgcgaagaatctgatttcac	723
QY	563	CAGACCTATGTGTGATGCTGTGATACGAGGTGGACAAATATGCCAGAACCTGGATTTTGTC	622
Db	724	aaactctgtctcttgcattgcattgcattgcttctgtggaanaagccccattacactgycacaac	783
QY	623	AACCTTATGCGCTCAGACTTCATGCTTCATGCTCTTGGAGAAAG-----TCACGGGACATTAAC	676
Db	784	agccctctgagcgaaggggtgctgcaggaacagagggccaaagctcctactacaatctggaat	843
QY	677	AGCCCCCTCTACAAAGGCAAGAGAGATGGTGTGACGACGACAGCCTCAACGTGTGATGCT	736
Db	844	gctgtggyggtactgtgatatataaggaaggaatgycatcaagaaagtgtgcatggaatcccc	903
QY	737	GCTGTGCAACATATGGCTGTGCAAAAGGGACCCCTGCCAGCAACTGATCTTGGATGCTT	796
Db	904	acatactgacacactctctcaacactgagcctc-tg--cagaacacacccgttgygggccccctgc	960
QY	797	ACCTAGGAGACGCTCCTTCACTAGCTGGCTCTCTATCAGACACCAAGATGGGGCCCCAGGC	856
Db	961	tctgacctgtgagctgctgtgagcccatcacagaagcttcaggctctccctgctctattagag	1020
QY	857	ACAGGCTGTGGACATCCAGGGCCCCCTTACCAAGAAAGGGATGCTGCGCTACTATGA	916
Db	1021	atctgcagcttctcttaaaaggagccaaagatcaacagcgctcacaagatcagaagttccctac	1080
QY	917	GCTTCC--TCTTGAAGGGGGCCACCAACAGAGATCCAGAGATCAGAGAGTGGCCCTAC	973
Db	1081	gcagctcaagggggaacacagctggtgtggtctatgtatgtgaagatcaggaacccaaggt	1140
QY	974	ATCTTCGCGGAACACCAAGTGGGCTTGTATGATGTGTGAAGCTCAAAACCAAGTCC	1033
Db	1141	cagttcttaagaatlttaaacctgtggaagagccatgatctgtgtctattgacatgagatgc	1200
QY	1034	AGCTATCTGAAACCAAGAGGACTGGGGGGGCGCATAGTGTGTGGGCACTGAGCTTAGATGAC	1093
Db	1201	ttcaatgtgcaatctctgcacacaggggccccttaccctctgtgtca	1244
QY	1094	TTTGCGGCTTCTCTCTGCAACCAAGGCGCATACCCCTCATCA	1137
RESULT	6		
LOCUS	SSGP38KD	1733 bp	RNA
DEFINITION	S.scrofa	38kDa heparin-binding glycoprotein.	MAM
ACCESSION	247803		
MID	9634097		
KEYWORDS	gp38k; heparin-binding glycoprotein.		
SOURCE			
ORGANISM	Sus scrofa		
REFERENCE	Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Artiodactyla; Suiformes; Suidae; Sus.		
AUTHORS	1 (bases 1 to 1733)		
TITLE	Millis,A.J.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (19-JAN-1995)		
AUTHORS	Millis A. J., University at Albany,		
TITLE	Biological Sciences, 1400 Washington Avenue, Albany, NY, USA,		
JOURNAL	2 (bases 1 to 1733)		
REFERENCE	Shackelton,L.M., Mann,D.M. and Millis,A.J.		
AUTHORS	Identification of a 38-kDa heparin-binding glycoprotein (gp38k) in		
TITLE	differentialing vascular smooth muscle cells as a member of a group		
JOURNAL	of proteins associated with tissue remodeling		
MEDLINE	J. Biol. Chem. 270 (22), 13076-13083 (1995)		
FEATURES	95286589		
SOURCE	Location/Qualifiers		
	1..1733		
	/organism="Sus scrofa"		
	/clone="pBS38k"		
	/dev_stage="adult"		
	/tissue_type="smooth muscle"		

Query Match	Best Local Similarity	62.5%	Score 156;	DB 48;	Length 1733;
Matches 688;	Conservative 0;	Mismatches 388;	Indels 24;	Gaps 9;	
<pre> /cdone.lib="lamda zap it" 67..1218 /codon_start=1 /product="38 kDa heparin-binding glycoprotein" /db_xref="PID:9634098" /translation="MGLRVNQTFVALVLLQSAAKLYCYTSMQYIEGSGSC AIDPFLCTHITTFPANTSNVETDLEMDVDTLDTLNLKRNPKLITLTVSGEM SQRSLKASMTQSRFTFKSPPEFLRTHGDDLMISGRDRKRLITLVSEKEM FVRALPGTERLLSLGSVAGKVAIDRGYDIAQISQHLDFISLTFDFGAMQTVL HSPLEFROGDASDRFNMAVSYALVGLGAPANKLVMGIFLPTGSRFTLASSTVD PASPGIIGRFTKEGILAYEICDFLQAGTVRRPGQOPATKGNOMVGDQDQD KNKAKYIKSRQLGAMWYMTLDDDFENFGCNLRPEPLSATIDVLAQA" </pre>					
BASE COUNT	408 a	507 c	463 g	355 t	
ORIGIN					
79	gtgctcagacaaggltttgttggccctgtgtgtgtgtccagacgtgtgtgtcctcaagctg 138				
14	gtggcctgggacagggttcattgtcctgtctgtatgtatcccatggggcctctgcgcgaaactg 73				
139	gtttgtactataccaagctgtgtctcagtaaccggaggatgtgtgagctgtctccagat 198				
74	gtctctactcttcaccacactggggcccatagacagacggggggagctgcctctctcccaag 133				
199	ggcatcggaccccttcctctgtcagccaatatctacagcttttgcacataagaacaat 258				
134	gacttggaccccgacttggcaccaccctctatgacccttctgtgtgcatatgaccacac 193				
259	gagattgcacaccttgagtggaatgattgtgacgtctctatgacacactggaacaacatcaag 318				
194	cagctgacgacctgactgagtggaatgagagagactcttctacagagacttcaatggccttaag 253				
319	aacaggaaccccaacctgaagaccctcctgtctgttggagatgtgaaacttggttctcaa 378				
254	aagatgattcccaagctgmaagccctgtttgaccatcgagagctgagatttgcgcactcag 313				
379	agatttccaataatgctcccaaacctcagatctgcagagcttcatcaagtctgtgtcca 438				
314	aagttcacagatattgtagccacggccacacacccgtgacacttctgttgcacattggccatc 373				
439	ccattcctcggaccacatggtcttgatcgtgactgtgacctagcctgtgactccctgtg--gc 496				
374	aggtttctggcccaatatacagctttgacggccttgaccttgactctggagatgccagagac 433				
497	----ggag-----agacaagcggcatctccacactagctcaagagatgaagct 543				
434	caggggacccctgcgcgtatgacagagagggcttccaaacccgtgtagacagacttggccaat 493				
544	gagtttgaagaggaaagccctg--cct--ggaacagaagcgtctcctgtctcagcggagcgtg 600				
494	gctttccagcaggaagcccaacaccttcaggggaagaaagcccttcttctgtagatgacgcgtt 553				
601	tctgtagggaaaggtcgccattgaaagggctatgacatcgcccaagatatctaacacctg 660				
554	ccagctgggacaccttatgttgatgtctggatgacaggttgacaaataatgccacgaacctg 613				
661	gacttcaacagacctttgacctatgacctcagcgtgtcctggcgccaaactacagagac 720				
614	gattttgtcaacattatggcctacgacttccattgaccttggggagaaagtcacgggacat 673				
721	cacagtcctccctgttcgcgaagccagggggatgtgcaagcttcgcacagatctagcaatgtcagac 760				
674	aacagccccctctacagagagcacaagagagatggtgcacacagccacacacgttgat 733				
781	tatcctgtgagctatgtgtcgtagcctgtggggcccccaagcaataagctgtgtgagtgatc 840				
734	gctgtgttgcaacagtgctgtcacaabaagggaccccttcacagcaagctgattccttggcatg 793				
841	cccaactttggagagacttcaacttggcctctcca--agacag--ac--gtggagaccccg 897				

Db	Query Match	Score	DB	Length
79	gtggctcagacaggtttctgtgcccctgtgtgtctctcagagctgtgtcgtcctacaagctg	9.5%	156	1733
	Best Local Similarity 62.5%		Prod. No.	6,59e-125
	Matches 688; Conservative 0; Mismatches 388; Indels 24; Gaps 9;			

QY	14	GTGGCCTGGGCAAGTTTCAATGCTTCGTGATGATGCCATGGGGCTCTGCTGCAAAACTG	73
Db	139	gttcgctactataccagcttgytctcagtaaccggagagtgatcgygagcttctccagat	198
QY	74	GTCGTACTACTACCAACACTGGGCCAGTACAGACAGGGGGAGCTGCGTCTCTGCCAAG	133
Db	199	gcacatgacccctctctctcgcaccacaatcctacagctttgccaacataagaacaat	258
QY	134	GACTTGGACCCCACTTTCACCCACCACTCATCTTACGCTTGGCTGACATGCACAAAC	193
Db	259	gagaattgacacacttggagtgaaatgtatgagctctatagacaactgaacactcaag	318
QY	134	CAGCTGAGACACACTGATGTGGAAATGACGAGACTCTTACCAAGAGTTCAATGGCTGAAG	253
Db	319	aacggaaaccccaacactgaaagaccctctgctgtgtgagatggaacttgytctcaa	378
QY	254	AAGATGAAATCCCAAGCTGAAGACCCGTGTATAGCATGSGAGGGTGGAATTTGGCACTGAG	313
Db	379	agatttccaaaataagcctccaaacactcagagtgcagacttcatcaatcggtgcc	438
QY	314	AAGTTCACAGATATGATGATGACGACGGCCACAAACCGCTCAACCTTTGTCAACTGGCCATC	373
Db	439	ccattccgcgagaccctatgcttctgtatgagctgagactagcttgcatacccccctg	496
QY	374	AGGTTTCTGGCCAATATACACTTTTGAACGGCTTGAACCTTGACTGGAGATACCCAGAAAC	433
Db	497	----ggag-----agacaagcgcgcatctcaccactttagtcaagagatgaagct	543
QY	434	CAGGGAGCCCTGCGCTATGACAAAGAGAGCGCTTACAAACCTGTATACGACTTTGGCCAT	493
Db	544	gagtttgaagggaaacccctg-cct--ggaacagagcgcgtctcgtcagcggagcgtg	600
QY	494	GCTTCCAGCAGAGAAAGCCAGACCTCAAGGAAAGAAAGGCTTCTTGTGATGCACGCTT	553
Db	601	tcgtgagaaagagcgcgcatttgcagaaggtatgacatgcgccagatatacaacactg	660
QY	554	CCAGCTGGGAGACCTATGTGTGATGCTGGATATACAGAGTGGACAAATGCCCCAGAACTG	613
Db	661	gacttcalcagccttlttgaactatlgatccacggtgcctgcgcgccaactacaagacac	720
QY	614	GATTTGTCAACCTTATGAGGCTACAGACTTCATGAGCTTGGGAGAAAGTTCAGGGAGAT	673
Db	721	cacagtcacctgttccgaagccaggggatatgcaagtltcgcagatltcagcaatgcagac	780
QY	674	AACACCCCCCTCTCAAGAGGCAAGAAAGACAGTGTGCAGACGACCTCTCAACGTGAT	733
Db	781	tatgtgtgagctatgtgctgagagctgctggggcccccagcaataagcttgtgtgtatc	840
QY	734	GCTGCTGTGCAACAGTGGCTGTCAAGAGGGAGACCCCTGCGCAACGATGATCTTGGCAGT	793
Db	841	cccaactttggagagagcttcaacttctgtgctcttcca-agacag-ac-gtggagcccg	897
QY	794	CCTACTACGGAGCCTCTTCAACCTTGGCTCTCTCATACAGACACAGATGGGGGCCCA	853
Db	898	gcattagggccagaataccaagccggtltcaacaagagaaaggtatcctgtcctactat	957
QY	854	GCCAAAGGGTCTGGCACTCCAGGCCCCCTTACCAAGAAAGAGAGGATGCTGGCTACTAT	913
Db	958	gagatctgtgacttccttcccaagagccacagtcctgtagacccttbgcagcagtgccc	1011
QY	914	GAAAGCTCT--CTCTGGAGAGGGGGCCACCAAAACAAGATCCAGGATATGAAAGTGGCC	970
Db	1018	tatgcacacaagggcaacacagttggtgtggtatgacatgcagaagagcgttcaaaaaaag	1077
QY	971	TACATCTTCGGGACAAACAGTGGGTGGCTTTGATGATGTGTGAGAGCTTCAAAACAAAG	1033
Db	1078	gcaaaagtacctgaagagcagcagctgctgctgtgtatagtgtgtgaccccttgagctgag	1137
QY	1031	GTCACACTATCTGAAGACAAAGAGCACTGGCGGGGCCATGTGTCGGCACTGGACTTATAG	1090
Db	1138	gaacttcggggcgaactctg 1157	
QY	1091	GACTTGGCGGCTTCTCTCTG 1110	





REFERENCE 1 (bases 1 to 1616)  
 AUTHORS Morrison, B.W. and Leder, P.  
 TITLE neu and ras initiate murine mammary tumors that share genetic markers generally absent in c-myc and int-2-initiated tumors  
 JOURNAL Oncogene 9 (12), 3417-3426 (1994)  
 MEDLINE 95060797  
 REFERENCE 2 (bases 1 to 1616)  
 AUTHORS Morrison, B.W.  
 TITLE Direct Submission  
 JOURNAL Submitted (14-NOV-1995) B.W. Morrison, Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA  
 FEATURES  
 source Location/Qualifiers  
 1..1616  
 /organism="Mus musculus"  
 /strain="FVB/N"  
 /dex\_stage="adult"  
 /tissue\_type="breast"  
 /cell\_line="SMF murine breast tumor"  
 /clone\_id="lambda ZAP II"  
 66..128  
 /gene="brp39"  
 /note="leader sequence"  
 66..121  
 /gene="brp39"  
 /note="homology to glycosyl hydrolases"  
 /product="BRP39 protein"  
 /db\_xref="PID:e211795"  
 /db\_xref="PID:91085066"  
 /translation="MGMRALJTGFAVLMILQSCASAKLYCFTSMQYREGVSFLPD  
 AIOPLCTHIYSPANISNDMLSTWENDESDNKLKTRNTNKLTLISVGKMF  
 GEKFESEIANTERTAFVSAPFLRSBGDGLAMLYRRLDKOYFSTLILKELNA  
 EFTENOPREKLLISAISAGKAIDTGYDIAOIAOHIDINATYDPRHWQITG  
 HHSLEFGQKDRDRSNVYAVQIMLRGAQSKLMLGIPITGKSTLASSENQLS  
 APISEGLPGRFTKEAGLAYEICDFLKEGVRLNENKVFPTKSNOMWGYEHKES  
 VKNRVGELKEKRLAGAMWALDDFDQGTCPKKEFFPLTNAIKDALA"  
 BASE COUNT 420 a 442 c 404 g 350 t  
 ORIGIN  
 Query Match 8.3%; Score 135; DB 81; Length 1616;  
 Best Local Similarity 62.5%; Pred. No. 3,23e-103;  
 Matches 674; Conservative 0; Mismatches 377; Indels 27; Gaps 13;

441 GCCCTGCCTAGACAGAGGCGCTTCAACACCTGTGTACAGGAGTGGCCATGCTTCC 500  
 Db 553 caaaggaggtccag--c-caggcagagagaacatccctgcagcagcgtttgacag 609  
 QY 501 AGCAGGAAGCCCAACCTCAGAGGAGAACGCCCTTCTTGAGTGCAGCGGTTCACACTG 560  
 Db 610 gaaagtgccatgacatggtcgtatgacatgcgcgcagatagcccaacatgatttta 669  
 QY 561 GGCAGACCTATGTGATGCTGTGATACGAGTGGAGCAAAATTCGCCCAAGACCTGATTTTG 620  
 Db 670 tcaatcctaatacctaagatttccatgagtcgtgcgcgaatacacaagccacacagcc 729  
 QY 621 TCAACTTATGCTTACGACTTTCATGCTGTGGAGAGGTGACCGGACATTAACGCC 680  
 Db 730 ccccttcccaagcccaagagacactaggtttgcacatacagacatgtaactatgcg 789  
 QY 681 CCTCTTAAAGAGGCAAGAAAGAGTGTGACGACGACGACCTTACGTGATGCTCTG 740  
 Db 790 tgcagtatatatacgtctctggagccagagccagcaagctatgagtgccatccacct 849  
 QY 741 TGCACAGTGTGCTCAGAGGAGGACCCCTGCCACAGCTGATCTTGCCATGCTTACT 800  
 Db 850 ttgggaagagcttcaactctgcatcttc-tgaa-atcagc-tggagctccaatcag 906  
 QY 801 ACGGACGCTCCTTCACTAGCTGCTCCATCATGACACACAGAGTGGGGCCACCCACAG 860  
 Db 907 gggaagagttccagcgcgtttcccaagagagcaggaacccctggcctaagatat 966  
 QY 861 GGTGTGGCACTCCAGGCGCCCTTACCAAGAGGAGGATGCTGCTACTATGATGATCT 920  
 Db 967 ggcagcttccaaaggagctgaatagatagctctcccaagagaagttccctgcta 1026  
 QY 921 GCTCTGCG--AAGGGGCCACCAACAGAAATTCAGATCAAGATGCGCTTACT 977  
 Db 1027 ccaagggaacagctgggtgggtatgagcaagaagagagtgtaaaacaagttggt 1086  
 QY 978 TCCGGGACAAACCAATGGGTGGCTTGTGATGTGAGAGGCTTAAACCAAGTACGCT 1037  
 Db 1087 tctcgaaggagaagaagctggcagagccatggtgtgtgagcagttgattt 1144  
 QY 1038 ATCTGAACACAGAGGAGCTGGCGGGCCATGCTGTGGCAGCTGACTGATGATCT 1095

RESULT 10  
 ID MND757 standard; RNA; ROD: 1506 BP.  
 AC D87737;  
 NI 91515818  
 DT 15-SEP-1996 (Rel. 49, Created)  
 DT 15-SEP-1996 (Rel. 49, last updated, Version 1)  
 DE House mouse; Musculus domesticus bone marrow cells mRNA for ECF-L precursor, complete cds.  
 KW ECF-L precursor.  
 OS Mus musculus (mouse)  
 OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;  
 OC Theria; Eutheria; Rodentia; Myomorpha; Muridae; Murinae.  
 RN [1]  
 RP 1-1506  
 RA Owashi M.;  
 RT Submitted (06-SEP-1996) to the EMBL/Genbank/DBD databases.  
 RL Makoto Owashi, The University of Tokushima, Faculty of Integrated  
 RL Arts and Sciences, 1-1 Minami-Johsanjima, Tokushima, Tokushima 770,  
 RL Japan (E-mail:Owashi@els.tokushima-u.ac.jp, Tel:0886-56-7261,  
 RL Fax:0886-56-7298)  
 RN [2]  
 RP 1-1506  
 RA Owashi M., Arita H., Taoka Y., Hayai N.;  
 RT "Molecular cloning and characterization of T lymphocyte-derived  
 RT eosinophil chemotactic factor (ECF-L)";  
 RL unpublished.  
 FH Key Location/Qualifiers  
 FT source 1..1506

```

FT      /organism="Mus musculus"
FT      /cell_type="bone marrow cells"
FT      /tissue_type="bone marrow"
FT      1..1197
FT      CDS
FT      /gene="MECF-L"
FT      /codon_start=1
FT      /product="ECF-L precursor"
FT      /db_xref="PID:g1545819"
FT      /translation="MAKILIVTGLAILNLVOLGSSYOLMCYTSMAKDRPIEGSPKGP
N      IDPCLCTHLIYAFAGMOMNEITYTHEQDRLDYELNGLKDNKTELKTLALIGMKRGP
A      SFSAMVSTQONQIIFQSVIRFLQYNFQGLNDMWPSSRSRSPPKDKLFSVLKEM
FT      KAFEESEVKDIPRLLLTSTGAGIIDIVIKSGYKIPELQSOLDYIQVMYDLDHPKDGX
FT      GENSPLYKSPYDIGKSDNLVNDISIYMKDHGAASEKLIYGFPAYGHTFILSDPSKTG
FT      GAPITISTGPGKRYTDESGLIAYEVCFLNEGATEVWDAPQEVPIAYOGNENWGDVNV
FT      SFLKRAQWLKDNNGAVVWRDMDDFSQSFCHQRHFPILTSLKGLDNLTHSASCKGPY
FT      "
SQ      Sequence 1506 bp; 422 A; 337 C; 335 G; 412 T; 0 other;

Query Match      7.5% Score 123; DB 20; Length 1506;
Best Local Similarity 57.2% Pred. No. 5,92e-91;
Matches 610; Conservative 0; Mismatches 451; Indels 6; Gaps 4;

Db 70 ctgattgctactataccagcttggtggaagagagcaatagaaggagttccaacct 129
QY 71 CTGCTGTCTGCTACTTTCACCACTGGGCCCATAGACAGAGGGAGCTCCTTCTGCCC 130
Db 130 gtaataattgaccctcgtgtgtactacacatgactatgctcttggtgaatgcaaat 189
QY 131 AAGCACTTGAGCCCAAGCTTTGACACCACCTCATCTACCTCGCTGGACGACCAAC 190
Db 190 aatgagataccttaacacatgagcaagactcgtgtacatataagcattgaatgctgt 249
QY 191 CACCAGCTGACACACACTGAGTGAAGACAGACACTCTTACACAGAGTTCATAGGCTG 250
Db 250 aaagaagaagaacatgagctgaacaaactctctgcccacttggaatggaatggaacct 309
QY 251 AAGAAAGTGAATCCCAAGCTGGAAGACCTCTTGCCATCGAGGCTGGAAATTCGGGACT 310
Db 310 gccctgtcagtgccatggtcttactcctcagaaccgctcagatatcatcagtcagtt 369
QY 311 CAGAACTTCAGATATGTAGACAGGCGCAACACCGTCAGACCTTGTCAACTGGGCC 370
Db 370 atcagattccttcgtcaataactttgatgctccaactcgtgactgagtcacctggg 429
QY 371 ATCAGGTTTCTGCGCAAAATCAGCTTTGACGGCTTGACCTTGAGTGGAGTACCCAGGA 430
Db 430 tctcggagggaagcctctcctaaggaacaacatctcttcagttcttggtgaaggaaatgcgt 489
QY 431 AGCCAGGGGAGCCCTGCGTAGACAAAGAGCGCTTCAACCTGGTAGAGACTTGGCC 490
Db 490 aaagcttttaagaagaatcgttgaagaagacattccaaggtcgtctaccacttcacaa 549
QY 491 AATGCCCTTCAGCAGAGAACGCCAGACTCTCAGGGAAGAAAGCGCTTCTGTGAGTGCAGCG 550
Db 550 ggaagcaggaatcatgacgtaatacagtcgtggtaacaagaatccctgaagatgctcagtt 609
QY 551 GTTCCAGCTGGGAGACCTATGTGATGTGATGATGAGTGAACAATAATCGCCAGAAC 610
Db 610 ctgactataatcagtcagacatatgatctcagatctcctaagatgagtcactacacgga 659
QY 611 CTGGATTTTTCACACCTTAAGCTTCAGACTTCATGCTCTTGGGGAAGGTCGACGGGA 670
Db 670 gaaataagtcctcctataatctcatalgacatggaagaagtcgtgactcaaatgtg 729
QY 671 CAATAACGCCCCCTCTACAAGAGGCAAGAGAGTGTGTACACAGCCACCTCAACGTG 730
Db 730 gattcaatcattctcactaggaagacattgagcagcttctagaagagctcattgtggga 789

```

```

QY 731 GATGCTGCTGTGCAACAGTGGCTGTGCAAGAGGAGCCCTGCCAGCAAGCTGATCTTGGC 790
Db 790 ttccagcatatgagcatacccttaccctgagtgacccttctaagactggaattgtgtcc 849
QY 791 ATGCCCTACCTAGGAGAGCTCCTTACACATGCGCTCCTCATACACACAGAGTGGGGGCC 850
Db 850 cctacaattagctctgcccacagaaagtacacagatgatacaggaactcctgtgttac 909
QY 851 CCAGCCACAGGGGTGTGCGACTCCAGCCCTTACACCAAGAGAGGAGTGTGCGCTTAC 910
Db 910 tatgagttgttcatattctgaaatgaagagcaccgaggtctgtggaatgcccccaagaa 969
QY 911 TATGAAGTCTG--C--TCTGGAAG--GG--GCCAACAACAGAAATCCAGATCCGAAGA 964
Db 970 gtaccctacgtcatcaaggttaatgagtggtgtgtgtatgacatgtcaagagctccaag 1029
QY 965 GTGCCCTACATCTTCGGGAGCAACAGTGGGTGGCTTTGATGATGTGAGAGACTTCAAA 1024
Db 1030 ttgaagctcagtggtctcaagacaacaaattagaagtggtcggtgtgtcgtcgtgac 1089
QY 1025 ACCAAGTCAAGTATCTGAAGCAGAGGAGTGGCGGGGCCCATGCTGTGGCACTGAC 1084
Db 1090 atgagatgactcagtggttcttctctgtcacccagagacattccctt 1136
QY 1085 TTGATGACTTGTGCGGCTTCTTCTGCAACAGGGCCGATACCCCT 1131

RESULT 11
LOCUS      D87757      1506 bp      mRNA      ROD      13-SEP-1996
DEFINITION House mouse; Musculus domesticus bone marrow cells for ECF-L precursor, complete cds.
ACCESSION  D87757
NID        91545818
KEYWORDS   ECF-L precursor.
SOURCE      Mus musculus bone marrow cells cDNA to mRNA.
ORGANISM   Eukaryota; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  1 (bases 1 to 1506)
AUTHORS   Owahshi,M.
TITLE      Direct Submision
JOURNAL    Submitted (06-SEP-1996) to the DBJ/EMBL/GenBank databases. Makoto Owahshi, The University of Tokushima, Faculty of Integrated Arts and Sciences; 1-1 Minami-Johsanjima, Tokushima 770, Japan (E-mail:owahshi@arts.tokushima-u.ac.jp, Tel:0886-56-7261, Fax:0886-56-7298)
REFERENCE  2 (bases 1 to 1506)
AUTHORS   Owahshi,M., Arita,H., Taoka,Y. and Hayaai,N.
TITLE      Molecular cloning and characterization of T lymphocyte-derived eosinophil chemotactic factor (ECF-L)
JOURNAL    Unpublished (1996)
FEATURES   location/Qualifiers
            source
            1..1506
            /organism="Mus musculus"
            /cell_type="bone marrow cells"
            /tissue_type="bone marrow"
            1..1197
            /gene="MECF-L"
            /codon_start=1
            /product="ECF-L precursor"
            /db_xref="PID:g1545819"
            /translation="MAKILIVTGLAILNLVOLGSSYOLMCYTSMAKDRPIEGSPKGP
            NIDPCLCTHLIYAFAGMOMNEITYTHEQDRLDYELNGLKDNKTELKTLALIGMKRGP
            PASFSAMVSTQONQIIFQSVIRFLQYNFQGLNDMWPSSRSRSPPKDKLFSVLKEM
            KAFEESEVKDIPRLLLTSTGAGIIDIVIKSGYKIPELQSOLDYIQVMYDLDHPK
            DGYTGENSEPLKSPYDIGKSDNLVNDISIYMKDHGAASEKLIYGFPAYGHTFILSDP
            SKTGAPITISTGPGKRYTDESGLIAYEVCFLNEGATEVWDAPQEVPIAYOGNENW
            GYDNVRSFLKRAQWLKDNNGAVVWRDMDDFSQSFCHQRHFPILTSLKGLDNLTHSA
            SCKGPY"

BASE COUNT  422 a      337 c      335 g      412 t
ORIGIN

```





Accession	Partial cds.
NID	U17988
KEYWORDS	g602293
SOURCE	
ORGANISM	sheep.
REFERENCE	Ovis aries
AUTHORS	Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
TITLE	Veterinaria; Euthera; Artiodactyla; Ruminantia; Pecora; Bovidae;
JOURNAL	Bovidae; Caprine; Ovis.
REFERENCE	1 (bases 1 to 1596)
AUTHORS	Marshall,J.T., Brownlee,A.G. and Nancarrow,C.D.
TITLE	Cloning and sequencing of a cDNA encoding ovine oestrous
JOURNAL	glycoprotein
REFERENCE	Reprod. Fert. Dev. 8 (2), 305-310 (1996)
AUTHORS	2 (bases 1 to 1596)
TITLE	Marshall,J.T.A.
JOURNAL	Direct Submission
REFERENCE	Submitted (02-DEC-1994) James T. A. Marshall, Division of Animal
AUTHORS	Production, Commonwealth Scientific and Industrial Research
TITLE	Organization, Locked Bag 1, Delivery Centre, Blacktown, New South
JOURNAL	Wales, 2148, Australia
FEATURES	Location/Qualifiers
SOURCE	1..1596
	/organism="Ovis aries"
	/strain="merino"
	/sex="female"
	/cell-type="epithelial cell"
	/tissue-type="ampulla oviduct"
	/dev_stage="oestrus"
	<1..1596
	/codon_start=1
	/product="92kd oestrus-associated oviduct glycoprotein"
	/db_xref="PID:g602294"
	/translation="LLVLKHKHDAHKLVCYPTNMAFSRPGASILPRDDPLCTH
	LVEFANNNNOIVKDPIDELKILPEPNKIKERNRGKATLISGVNMFSGRFTKMI
	TSFNREKFNVSVALRTDHFGEFDGLDFEPLGSRARBRWFLELLELQAFKK
	EAQITMRPRLLSAASGDPVIAOKAYARLGLDPLISLSDHSGWREKTYGHS
	PLESLPDPKRSATAMSTWRQLGVPEPELNLGPTGTHLRLASDLEAGVAPPP
	SPGRTKQAGFLAYEVCSFVQAKRKMINQYVYPAKKEWGYDAISFGYKAEF
	IKRHFGAMWATLDDLRNFCGTGEPFLVHTLNNLLVDEFSSTPSPFWSTAW
	NSSSIEPMPTMDLTGLGILPLGGEAVATEHRKSAWTTPRGETATPTPLSL
	SGRTAAPEGKTESGGERPLFVSGHLAVSPGIAVGPVHLIGQKVPDPKAGVPEK
	VTPSGKWTVTPDGRATLERRL"
	<1..>1596
	/product="92kd oestrus-associated oviduct glycoprotein"
	<1..36
	/note="inferred from bovine sequence: GenBank Accession
	Number D16639"
	/codon_start=1
	1..19
	/note="derived from PCR primer based on bovine CDNA
	sequence: GenBank Accession Number D16639"
	/note="derived from PCR primer based on bovine CDNA
	sequence: GenBank Accession Number D16639"
	BASE COUNT 352 a 440 c 430 g 374 t
ORIGIN	
Query Match	5.7%; Score 93; DB 47; Length 1596;
Best Local Similarity	58.7%; Pred. No. 6,99e-61;
Matches	364; Conservative 0; Mismatches 253; Indels 3; Gaps 2;
Db	40 aagctggtgtgtattattcaccacattgagcatcagtcggccggtctgctcgatctcg 99
Oy	68 AAACGTGCTGCTACTTCCACCAATGCGCCAGTACAGACAGGGGAGSGCTCGTCC 127
Db	100 ccccgagatctgagaccctttcttcgacacacacctggtatttgctcttcctcgatgaac 159
Oy	128 CCCAAGGACTTGAGCCCAAGCCCTTGACACCACTCATCTAGCCCTCCCTGGCATGAC 187
Db	160 aacaatcagatgttctcctaagatcccccgagtagaanaatccttaaccagatitcaac 219
Oy	168 AACCAACGAGCTGACGACCACTGAGTGAATGACGAGAC--TC-TTACCAAGAGTTCAAT 244

Db	220	aagctcaagggaagaacaggggagaaacaactgcgtcgttgggggttggaatttc	279
Oy	245	GGCTTGAAGAAGATGATGCCAAGCTTAACACCCTGTTAGCCATTGCGAGGCTTGAAATTTTC	304
Db	280	ggcacctcgaggtttaccagaatgctgtgccaggtttccaacccgggaaaagtgttccaat	339
Oy	305	GGCACCTCAGAAGTTTCACAGATATGTGTAGCCAGGCAACACCGTCAGACTTTTTCACAC	364
Db	340	tcaaagtgcgccctcccgtaggaacacatggttcgatgtgtctgtaacctttctcttgttac	399
Oy	365	TGGGCATCAGAGTTCTTGCCGAATTAAGACTTTGACGGCCTTACCTTGTACTGGAGTAC	424
Db	400	ccttgactcaagagcaagccccgcgcgtgacgccttgacactttctcttaacttgaagag	459
Oy	425	CCAGGAAGCCAGGGGAGACCTCGCCGTAGAACAGAGAGCGCTTCAACACCCCTGGTACAGAC	484
Db	460	ctctctgaagcgcttcaagaatlagcccacatacctatgcgcaccaagagctgtcgtctct	519
Oy	485	TTGGCCATATGCTTCCAGAGGAAGCCCCAGACCTCAGGGAAGGAAGCGCTTCTTGAGAT	544
Db	520	gcttcgctctcttgggagccccacgcatatccagaagcatatgtcacgccttcgggc	579
Oy	545	GCACGGTTCCAGCTGGGCAAGACCTATGTGTGAATGCTGATACAGAGGTGACAATAATCGCC	604
Db	580	agaactcctgatatcattcacagcgtcttcttcaagacttacaagcaagcgtbgaaaaagtc	639
Oy	605	CAGAACCTCGATTTTGTCAACCTTATGCGCTACGACTTCCATGCTCTTGGAGGAAGTTC	664
Db	640	acaggaacacaatagccctct	659
Oy	665	ACGGGACATATACAGCCCCCT	684

RESULT	15		
LOCUS	BABEDOSG	2228 bp	mRNA PRI 11-OCT-1996
DEFINITION	Papio hamadryas anubis estradiol-dependent oviduct-specific glycoprotein mRNA, partial cds.		
ACCESSION	M59903		
NID	g1256950		
KEYWORDS	estradiol-dependent oviduct-specific glycoprotein.		
SOURCE	Papio hamadryas anubis cDNA to mRNA.		
ORGANISM	Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Cercopitheidae; Cercopithecinae; Papio.		
REFERENCE	1 (bases 1 to 2228)		
AUTHORS	Donnelly,K.M., Fazleabas,A.T., Verhage,H.G., Mavrogianis,P.A. and Jaffe,R.C.		
TITLE	Cloning of a recombinant complementary DNA to a baboon (Papio anubis) estradiol-dependent oviduct-specific glycoprotein Mol. Endocrinol. 5 (3), 356-364 (1991)		
JOURNAL	91367180		
REFERENCE	2 (bases 1 to 2228)		
AUTHORS	Jaffe,R.C., Arias,E.B., O'Day-Bowman,M.B., Donnelly,K.M., Mavrogianis,P.A. and Verhage,H.G.		
TITLE	Regional Distribution and Hormonal Control of Estrogen-Dependent Oviduct Specific Glycoprotein mRNA in the Baboon (Papio anubis) Biol. Reprod. (1996) In press		
JOURNAL	3 (bases 1 to 2228)		
REFERENCE	Jaffe,R.C.		
AUTHORS	Direct Submission Submitted (05-MAR-1991) Randal C. Jaffe, Department of Physiology and Biophysics, University of Illinois College of Medicine, Chicago, IL 60612-7347, USA		
FEATURES	Location/Qualifiers		
SOURCE	1..2228 /organism="Papio hamadryas anubis"		
CDS	13..1884 /codon_start=1 /product="estradiol-dependent oviduct-specific glycoprotein" /db_xref="PID:g1256951"		

BASE COUNT 559 a 572 c 531 g 566 t  
ORIGIN

Query Match		5.7%	Score 93:	DB 67	Length 2228:
Best Local Similarity		58.7%	Pred. No.	6,99e-61:	
Matches	364;	Conservative	0;	Mismatches 253;	Indels 3; Gaps
Db	79	aaactggtgttatttaccacacttggcacacagtcgcccagcccgctcagatctg	138		
Oy	68	AAACGGGTCTCTACTTCCACCACTGGGCGCCAGTACAGACAGGGGGAGTCGCTCTG	127		
Db	139	ccccatgcccggcccccttctctgtgaaaccaactgatattgcttgcctcaatgaac	138		
Oy	128	CCCAAGACTTGGACCCAGCCTTTGGACCCCACTCACTTACGCTTTCGCTGGCATGAC	137		
Db	199	aacaatcagatttgttctaagaatctccaaaggacagaaatctctacccagaattcaac	258		
Oy	188	AACCAACCAGCTGAGACCACTGAGNGAATGAGAGAC--TCTCTACAGAGGTTCAAT	244		
Db	259	aaactaaaggaggagaaagagaactgaaaacactgtcttccatcgycgggtggaaactt	318		
Oy	245	GGCTGAAAGAGATGAATCCCAACACTGAAGACCTGTTAGCCATCGAGGCTGGAAATTTC	304		
Db	319	ggcacctccagattccaccactatgttgccacatttgcacaacgctgaagaatttatgtc	378		
Oy	305	GGCACTCAGAAATTCCACGATATGTGTAGCCAGGCGCAACAAACGTTAGACTTTGTCAAC	364		
Db	379	tcagttataatccctcttgagagacaatgaacttgaatgtgtcttcttttctctatat	438		
Oy	365	TCGGCCATCAGAGTTCTCGCAATAATACAGCTTTTACGGCCTTGACCTTGACTGGAGTAC	424		
Db	439	ccctgaactaagagcagccccatgatcatgcagcgtgagacttctctcttaattgaagag	498		
Oy	425	CCAGGAACCCAGGGGAGGCCCTCCGGTAGACAAAGAGCGCTTCAACAACTCTGTACAGAC	464		
Db	499	ctctcgttttccttcgcgaagagaggaactgtctcaacatcgcccagaagctgtcgtct	558		
Oy	485	TTGGCCAAATGCCCTTCCAGACGAAGAGCCAGACCTCAGGGAAGAAACCCCTTCTCTAGT	544		
Db	559	gctgtgtttcttgggtgtccacaatatgltccaacaatcctatgtgtgycgttcttga	618		
Oy	545	GAAGGGTTCAGCTGGGCGACACCTATGTGGATCTGGATTCGAGGTGACAAATGCGC	604		
Db	619	agactccttggatttcaatcatgtcttctctatgaactacatgaagattggaaaaagt	678		
Oy	605	CAGAACCTGTGATTTTGTCAACTTATGGCCTTAGCACTTCATGCTCTTGGAGAAAGCTC	664		
Db	679	acagagactaatagcccccct	698		
Oy	665	ACGGACATTAACAGGCCCT	684		

Search completed: Sun Jun 29 19:55:25 1997  
Job time : 1542 secs.







```
QY      84 TCACCAACTGGGCCAAGTACAGACAGGGGAGGCTCGCTTCTGCCCCAAGGACTTGGACC 143
Db      214 gcttcctgtgtaccatcatcatcagcttgcacaaataagaacaagcatcacatcgaca 273
QY      144 CCAAGCCTTTGACACCACTATCTACGCCCTTCCGTGGCATGACCAACACACAGCTAGACA 203
Db      274 cctggagatggaaatgtatgtacgctctacagcaltgtctcaaacactcaacaacagaaac 333
QY      204 CCACTGAGTGAATGACGAGACTCTCTACAGAGATTCAATGGCCCTGAAGAAGATGATAC 263
Db      334 ccaacttgaagaactctctctgtctgtcgaggaatggaaacttggcttccaagaatttcca 393
QY      264 CCAAGCTGAAGACCTGTTGATCCATCGGAGCTGGAATTTCCGACACTCAAGAATGTCACAG 323
Db      394 agatagctctcaaacacagagctgcgcggaacttcacatgaagtaacgcacattctgc 453
QY      324 ATATGGTAGCCAGGCCCAACACCGTACAGCTTTGCAACTGGCCATAGAGTTTCTGC 383
Db      454 gcaacccatggcttgaatggcgctgaacctgtgcgtctacacctgga---c---ggag-- 505
QY      384 GCAAAATACAGCTTTGACGGCCTTGACCTTGACTGGAGATACCAGAAAGCCAGGGGAGGCC 443
Db      506 -----agcaacaacacatttaccacccctaatcaaggaatgaagccggaatttata 558
QY      444 CTGCCGTAGACAAGAGAGCGCTTACAAACCGTGTACAGACTTGGCCAAATGCCCTTCAGC 503
Db      559 aggaagcccgag--c-cagggaaaaaagcagctcctgcctcagcgacgcactgtctgcggga 615
QY      504 AGGAAGCCCAAGACTCTCAGGGAAGAAAGCCTTCTTGATGTGACGGCGTTCCAGATGAGC 563
Db      616 aggtcaccattgacagacgtatgacatgtacatgcaagaatatacccaacccgtgatttca 675
QY      564 AGACCTATGTGATGCTGATACGAGAGTGAACAATAATGCCCAAGAACTGGATTGTGTCA 623
Db      676 gcatcatgaactcagatatttcaatgagcgccgtggtgagacacagcgatcacaagctccc 725
QY      624 ACCTTAATGGCTTACGAGACTTCCATGGCTTGTGGAGAAAGTCCAGGGGACATACAGCC 683
Db      736 tgttcggaagtcaggaagatgtcaagctctgaacagattcagcaaacactgaactatgtg 795
QY      684 TCTACAAAGAGGCAAGAAAGATGCTGACAGCAGCCAGCTCMAAGTGAATGCTGCTTGC 743
Db      796 ggtacatgttgaagctgtgagctctctgcagtaagctgtgatatggatcccaacttcg 855
QY      744 AACAGTGGCTGCAAGAGAGGAGACCCCTGCAAGAGTGAATGCTTGATGCTATGCTACG 803
Db      856 gtaggagactcactctgtgctcttc--tgagactg--gtgttcacagcgcaatctcaagac 912
QY      804 GAGGCTCTTTCACACTGGCTCTCTCATACACACAGATGGGGGCCCCAGCCACAGGGT 863
Db      913 cgggaattcgaagccggtttacaaaggaagcaggaacctgtgacctactatagaatctgtg 972
QY      864 CTGGCACTCCAGGCCCTTACACCAAGAAAGAGGAGATGCTGCGCTTATTAAGTCTGCT 923
Db      973 acttctctcgcgagccacagatccatagaacctcgagcagcagtgctccatgcccaca 1032
QY      924 CCTGGAGG-G-GG-GCCACCAAAACAGAGAAATCCAGAGATCAGAAGTGCTCCCTACATCT 980
Db      1033 agggcaaacagctgtgtgtgatacgaacgaagaagcgtccaagaagaagtgcaatc 1092
QY      981 GGGACAACCAAGTGGTGGCTTTGATGATGTGAGAGCTTCAAAAACCAAGTCAAGCTATC 1040
Db      1093 tgaagagatgaagcagctgtgcagcgcccatgtatagcgcccggaactggaactccag 1152
QY      1041 TGAAGCAGAAGGAGACTGGGGGGGCCATGCTGTGGCACTGGAATTGATGACTTTGCCG 1100
Db      1153 gctcctctctgcggcgag 1170
QY      1101 GCTTCTCTCTCAACACAG 1118
```

RESULT 2  
ID N81756 standard; DNA; 966 BP.

```
AC      N81756;
DT      19-OCT-1990 (first entry)
DE      Gene encoding polypeptide involved in protective mechanisms
KW      Immune response; cell growth; ss.
FH      Key
FT      CDS
FT      /tag= a
PN      J63032898-A.
PD      30-JAN-1988.
PF      16-JUL-1986; 167518.
PR      16-JUL-1986; JP-167518.
PA      (DAIN) Daiinpon Pharm KK.
DR      WPI; 88-068419/10.
PT      P-PSDB; P81342.
PS      New polypeptide and DNA encoding it - related to protective mechanisms
PS      such as immune response etc.
CC      Disclosure; P; Japanese.
CC      Polypeptide encoded is involved in protective mechanisms such as immune
CC      response, cell growth and activation of protective functions.
SQ      Sequence 966 BP; 238 A; 282 C; 247 G; 199 T;
```

Query Match 8.1%; Score 133; DB 1; Length 966;  
Best Local Similarity 66.3%; Pred. No. 4,48e-71;  
Matches 270; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

```
Db      23 cagcttctgtgtaccatcatcagcttgcacaaataagaacaacgctcgtact 82
QY      24 CAGGTTTCATGTCGTCGCTGATGATCCATGGGGCTCTGCAAAACGTCTGCTACT 83
Db      83 acacacgctgtgtccagtaacccggaaagcgatggagctgtcctccagatgccctgacc 142
QY      84 TCACCAACTGGGCCAAGTACAGACAGGGGAGGCTCGCTTCTGCCCCAAGGACTTGGACC 143
Db      143 gcttcctgtgtaccatcatcagcttgcacaaataagaacaacgctcagatgaca 202
QY      144 CCAAGCTTTGACACCACTATCTACGCTTCCCTGCGATGACCAACACACAGCTGAGCA 203
Db      203 cctggagatggaaatgtatgtacgctctcagcagctgtctcaaacacacgaagaagaacc 262
QY      204 CCACTGATGATGACAGAGACTCTCTACAGAGATTTCAATGGCTTGAAGAAGATGAATC 263
Db      263 ccaacttgaagactctctctgtctgtcgtgagatggaaacttgggttcacaaagtcca 322
QY      264 CCAAGCTGAAGACCCGTGTTAGCCATCGAGGCTGGAATTTGGCACTCAGAATTCACAG 323
Db      323 agatagctcccaacaccagagctgcgcggaacttcaatgaatgaatgacgacattctgc 382
QY      324 ATATGGTAGCCAGGGCCACACACCGTCAAGACCTTGTCAACTGGGCCATCAGGTTCTGC 383
Db      383 gcaacccatggcttgaatgggtgagctctgacctgacctgaccttaccctgga 429
QY      384 GCAAAATACAGCTTTGACGGCCTTGACCTTGACTGGAAGTACCAAGCA 430
```

RESULT 3  
ID Q90443 standard; cDNA to mRNA; 2504 BP.  
AC Q90443;  
DT 10-JAN-1996 (first entry)  
DE Murine oviduct specific glycoprotein cDNA.  
KW Murine oviduct specific glycoprotein; recombinant production;  
KW MOGP; ds.  
OS Mus musculus.  
FH Key  
FT CDS  
FT /tag= a  
FT /note= "START codon absent"  
FT sig\_peptide 3..56  
FT /tag= b  
FT mat\_peptide 57..2156  
FT /tag= c  
PN J07107979-A.  
PD 25-APR-1995.  
PF 15-AUG-1994; 214227.



QY 392 AGCTTGAAGGCGCTTGAAGCTTGAAGTACCCAGGAGAGCCAGGAGGCC-TGCCGT 450  
Db 363 rrvnkmgrryhytgvnwmkndrntdnywmgdngsdnmaahysganknnwtg 422  
QY 451 AGACAAGAGACC-GCTTCAACACCTGTGACAGACTTGCGCAATCCCTTCCAGCAGAG 509  
Db 423 rrrnn-wv-kgsnndnnandndndscdklnstnanvngtntmnygsnnnrk 480  
QY 510 CCCAGACCTCAGGAGAGAGAGCCCTTCTCTAGTCAGAGCGTTCCAGCTGGCAGAGACT 569  
Db 481 nmnnknaasmwrrwnnnngnsnrhykgsagrsntnsrsgsysnmtahykynnaat 540  
QY 570 ATGTGATGTC-TGGATAGAGGTGGACAAATCCGCCAGAACCTTGATTTGTCAACTT 628  
Db 541 ghnkgnvnnkhnvknrrnntrnnnnkhmrdivnnhntnngschnnnnnrvnrycn 600  
QY 629 -ATGGCCTTACGACTTCCATGGCTCTTGGAGAGAGTCAAGGACATAACAGCCCTCTTA 687  
Db 601 rgsnndnnndnnndnnmrysnndnvkymannhnsnshysnksnscvvsrnnkn 660  
QY 688 CAAGAGGCAAGAGAGAGAGTGGTGCAGAGCCGCTCAACGTGGATGCTGTGTCAACA 747  
Db 661 tdynasnrstanndnnnanyaknnntannnnsgnnnttgmaadvyssngnnnnanr 720  
QY 748 GTGGCTGCAGAGAGGAGCGCCCTGCGCAGAGCTGATCCTTGGCATGCTTACAGGAG 807  
Db 721 sgnyngnngdnshknvknvknngnrynnrdtnnnnnnnnnnnnnnnnnnnnnnn 780  
QY 808 CTCCCTTCACTACCTGCTCTCTATCAGACACAGAGTGGGGGCCCAAGCCAGAGGTCTGG 867  
Db 781 nkgnnrnnknngtsnndnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn 840  
QY 868 CACTCCAGGCCCTTACACCAAGAGAGAGAGTGGTGGCTCTCTATGAAGTGTCTCTG 927  
Db 841 svannknngtynnanndsvtnynsdvngntansnntmmvtnndndytcnd--and 898  
QY 928 GAAGGGGCGCCACCAACAGAGATCCAGAGATCAGAGAGTGGCTCTCACTTCCGGGACA 987  
Db 899 ndvykvtngdagmvsngngngnrrhannanndavsnrrhrhndnnrrn 958  
QY 988 CCAGTGGGTGGCTTGTGATGTGAGAGCTTCAAAAACCAAGGTCAGCTATCTGAACA 1047  
Db 959 gvhtnvcagvnnkmycnngdvtvnlasrnnsgnanknhv 1001  
QY 1048 GAAGGAGCTGGGCGG-GGCCATGATC-TGGGCATCTGAGACTTG 1088

RESULT 5  
ID Q90442 standard; cDNA to mRNA; 1994 BP.  
AC Q90442;  
DT 10-JAN-1996 (first entry)  
DE Bovine oviduct specific glycoprotein cDNA.  
KW Bovine oviduct specific glycoprotein; recombinant production;  
KW BOP; ds.  
OS Bos taurus.  
FH Key Location/Qualifiers  
FT CDS 2..1615  
FT /tag- a  
FT /note- "START codon absent"  
FT sig\_peptide 2..55  
FT /tag- b  
FT mat\_peptide 56..1612  
FT /tag- c  
FT J07107979-A.  
PD 25-APR-1995.  
PF 15-AUG-1994; 214227.  
PR 19-AUG-1993; JP-227881.  
PA (KINO-) KINOSEI PEPTIDE KENKUYUSHO KK.  
DR MPI: 95-190179/25.  
P-PSDB: R73991.  
FT New DNA encoding an oviduct-specific glycoprotein - useful for  
FT recombinant protein production in high quantities.  
PS Claim 1; Pages 7-9; 22pp; Japanese.

CC Q90442 encodes R73991 bovine oviduct specific glycoprotein (BOP).  
CC The cDNA can be used for the commercial recombinant prodn. of  
CC BOP in high quantities.  
SQ Sequence 1994 BP; 445 A; 536 C; 520 G; 493 T;  
Query Match 5.2%; Score 85; DB 15; Length 1994;  
Best Local Similarity 58.1%; Pred. No. 4,11e-38;  
Matches 360; Conservative 0; Mismatches 257; Indels 3; Gaps 2;  
Db 59 aagctgtgttatttcccaacttggcattcagtcggccggccctgcctcaatcctg 118  
QY 68 AAACGTGCTCTACTTACCAACTGGCCCAAGTACAGACAGGGGAGGCTCGCTCTCTG 127  
Db 119 cctggagatcggcccggttctctgcagcccaactglatgttcttgcctgtagagc 178  
QY 128 CCAGAGACTTGGACCCAGCGCTTGGACCCACCTCATCTAGGCTTGGCTGGCATGACC 187  
Db 179 aacaatcagattgttcctaaagatcccaagatgagaataatcctctcccaagattgaac 238  
QY 188 AACCAACAGCTGAGACACACACAGTGAATGATGATGACAGAC--TC-TCTACAGGAATTCAAT 244  
Db 239 aagctcaaggaggaagacagagggctgaacagctgtctgtccatcgggggtggaactt 298  
QY 245 GGCGTGAAGAGATGAATCCCAAGCTGAAGACCTGTAGGCATCGGAGGCTGGAATTTC 304  
Db 299 ggcacggtgaggttcacacagatgtctccagctttcccaacggggaaggttcgaagt 358  
QY 305 GGCACTCAGAAATTACAGATATGTATGTAGCCAGGCCAACAACCTTGATGTGTCAAC 364  
Db 359 tcaatgtagccctcctcctagagacacatggtcttgatgtctgagaccctctctctgtac 418  
QY 365 TCGGCATCAGATTCTTCGCAATATACAGCTTTTACGCGCTTGACCTTGAGTGGAGATAC 424  
Db 419 ccttgactagaaggcagccctgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgc 478  
QY 425 CCAGAGAGCCAGGGGAGGCGCTCGGTAGACAGAGAGGCGCTTCAACAACCTTGATCAAGAC 484  
Db 479 ctccctgagcgcttcaagaatgagagccagatccatgctccaaaggtcgtctgtct 538  
QY 485 TTGGCAATGCTTTCACACAGAGAGCCAGCTCAGAGGAAGAACCCCTTCTTCAAGT 544  
Db 539 gctgcgctctctgggagaccacatgctgcagaaagcgtatgaagacgctctctgggc 598  
QY 545 GCAGGGGTTCACAGCTGGGCAACCTATGTGATCTGTGATCGAGTGGAGCAAAATGCC 604  
Db 599 agactcctgattcatcagtgctgtcttcttgaacttaactgaagcctgggaaaggtc 658  
QY 605 CAGAACCTGGATTGTCACACCTTATGGCTTACGATCTCATGCTGTGGAGAGAGTGC 664  
Db 659 acagagacacaatagccctct 678  
QY 665 ACGGACATTAACAGCCCTCT 684

RESULT 6  
ID Q10572 standard; DNA; 1047 BP.  
AC Q10572;  
DT 09-APR-1991 (first entry)  
DE Human Natriuretic Peptide Receptor B.  
KW NPB; ANP; BNP; CNP; Kidney failure; heart failure; protein kinase;  
KW hyperaldosteronism; glaucoma; guanyl cyclase.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT Peptide 1..22  
FT /label- signal sequence  
FT Protein 12  
FT /label- mature NPB  
FT Domain 23..455  
FT /label- extracellular domain  
FT /note- "binds natriuretic peptides A, B and C"  
FT Domain 456..456  
FT /label- transmembrane domain  
FT Domain 479..1047

```

FT /label= cytoplasmic domain
FT /note= "GC and protien kinase activity"
FT Modified -site 24..26
FT /label= N-glycos_-site
FT Modified -site 35..37
FT /label= N-glycos_-site
FT Modified -site 161..163
FT /label= N-glycos_-site
FT Modified -site 195..197
FT /label= N-glycos_-site
FT Modified -site 244..246
FT /label= N-glycos_-site
FT Modified -site 277..279
FT /label= N-glycos_-site
FT Modified -site 349..351
FT /label= N-glycos_-site
FT Modified -site 600..602
FT /label= N-glycos_-site
PN MO9100292-A.
PD 10-JAN-1991.
PF 22-JUN-1990; 003586.
PR 23-JUN-1989; 05-370673.
PA (GETH ) GENENTECH INC.
PI Chang M, Goeddel D, Lowe D;
PI WPI: 91-036711/05.
DR N-PSDB; Q10324.
PT Natriuretic protein receptor B - for diagnosis and treatment of
PT kidney failure, heart failure, hyperaldosteronism, glaucoma etc.
PS Claim 3, Fig 1; 49pp; English.
CC The sequence was derived from the DNA encoding natriuretic peptide
CC receptor B. NPRB, having guanyl cyclase (GC) activity and protein
CC kinase activity. The DNA can be inserted into expression vectors
CC for the prodn. of the protein, opt. after being mutated to produce
CC NPRB analogues. The protein has a mol wt. of 115 kD (calculated Mr=
CC 114,952). The protein (or variants) can be used in treatment of
CC natriuretic peptide disorders, and also to isolate peptides using
CC affinity chromatography. Antibodies with affinity for NPRB can
CC also be prepd.
SQ Sequence 1047 BP; 87 A; 15 C; 83 G; 51 T;

Query Match 4.88; Score 78; DB 2; Length 1047;
Best Local Similarity 8.98; Pred. No. 1,92e-33;
Matches 83; Conservative 251; Mismatches 596; Indels 7; Gaps 7;

Db 28 avnnhnhnsyawnrvygnavaavnanvgrannvndrvnsnngnacsnnannasavdn 87
Cp 1623 AAGAGGGGCGCAAAACCAAGATTATTTTCAGTGAAAGGGCGACGCCAGACCCAG 1564
Db 88 kghdndnnngngcyvnaasvarnaashwrnnmtagavagsnaskndhyrtvrt-gnsa 146
Cp 1563 AAAAAAGGAAGGCGAGCGCTGAGAGCAAGACCTGTGATAAGAGACACAGAAAGGCC 1504
Db 147 nkngrnvntrhghnwtarannndardtrnhntnngnannngsnsvghnyarn 206
Cp 1503 TGCAGAGACCAAGTTGCGGGCCCCAGGAGAAACCCAGAGGCGCTGTAGATGATCCT 1444
Db 207 nggnnathnrranrgrnrvnncgnnnmhnannnannrntngdyvnyndvngsnragn 266
Cp 1443 GGGCCAGCCTCAAA-AGCTGGAGCTGAGAGGGCTTACGACATCAATTCAGGTGACGA 1385
Db 267 trrtgrnwdntrrmananrnanrnnvntlyrnnnnnnnnnnnnnnnnnnnnnnnnnnnn 326
Cp 1384 TT-TGCAGAGGTTGCTGAACACACAGGCGCTTCGGGAGCTTCTGTGAAACGCGCCCG 1326
Db 327 nsnmnnnagncvqnnnyannvnnnnnnngtrndgrnrvnkmgrryhytgvvmdkn 386
Cp 1325 CTCGACAGCTGTGAGAGCTGAGACCTCCCGAGAGATTGGATAGAGCCATTCAGCTTTGC 1266
Db 387 drntdvnwamgdndsgdnnaahysganknwvtrfnnnnwkygannsdnnncandnddn 446
Cp 1265 CCTGGAGAGAACTGTCTGTTCAGGCGCTGGGCCATGCTCAGTTCAGAGGCGTGCAGCTG 1206
Db 447 scdtknnsnranvngtgnnnmngvssnnnnnkmmnknnasnmwvnrwnnnnnngnsnr 506

```

```

Cp 1205 GTTTGGAACTTCAAGCTCTGGGGTCCCTGAAAGGCAAGATAGAAACATCACTCTCTGCC 1146
Db 507 yhkqagrsrntusnrgssysnmtahgkynnannatghknvannkhnkrnnntrnvn 566
Cp 1145 GTAGCGCTCGAGATGAGGGGGTATCGGCCCTGGTTGACAGAGAAAGCCGAAATCATCTCA 1086
Db 567 nkhmrtdvnnhntnngnacdnnnnnenvtynrnsndnnnnndnnndmnnrnsnnd 626
Cp 1085 AGTCAGTGCCGAGACCATGGCCCCCGCCAGTCCTTCTGCTTCACATAGCTGACCTTGG 1026
Db 627 nvkgnmnnhnsnshgnsncvdsrnnvnhntdygnasnrstannndnanyaknn 686
Cp 1025 TTTTGAAGCTTCCACATATCAAAAGCCACCATGCTTGTCCCGGAAGAT-GTAGGGC 967
Db 687 tannnsgnnnltgmaadvysngnnnnnnnansngnyngnngdsnkhnvkvrrgnr 746
Cp 966 ACCTTGATMCTCGATTCCTGTTGTGGGCCCTTCACAGAGACGACTCATAGTAG 907
Db 747 -ynrnsndtrnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn 805
Cp 906 GCCAGCATCCCTCTCTCTTCTGTGGAGAGGGGCGCTGGAGTCCAGACCTGTGGCTGGGGCC 847
Db 806 nuyannnnknnvnrtnaynnkrkanannnnnnnnhsvannnnkrngtvnannndsrvnys 865
Cp 846 CCCACTCTGCTGTCTATAGAGAGGCCAGTGTGAAGAGCGTCCGTAGTAGAGATGC-C 788
Db 866 dnvgntansansltmnnvlnndnlycndandndvkvntngdymvssngnrgnrg 925
Cp 787 AAGGATCAGCT-TGCTGGCAGGGGTCCTCTTGACGACACGTTCACAGACGATTCGA 729
Db 926 hannarnannandavsnrrnrhndnnrnnrgvht 962
Cp 728 CGTTGAGCTGGCTGCTGCACACCTCTCTTGTGCT 692

RESULT 7
ID 090444 standard; cDNA to mRNA; 2366 BP.
AC 090444;
DT 10-JAN-1996 (first entry)
DE Hamster oviduct specific glycoprotein cDNA.
KW Hamster oviduct specific glycoprotein; recombinant production;
KW HOGP; ds.
OS Cricetus griseus.
FH Key Location/Qualifiers
FT CDS 3..2009
FT /tag= a
FT /note= "START codon absent"
FT sig_peptide 3..36
FT /tag= b
FT mat_peptide 57..2006
FT /tag= c
FT PN 307107979-A.
PD 25-APR-1995.
PF 15-AUG-1994; 214227.
PR 19-AUG-1993; JP-227881.
PA (KINO-) KINOSBI PEPTIDE KENKYUSHO KK.
DR WPI: 95-190179/25.
DR P-PSDB: R73993.
PT New DNA encoding an oviduct-specific glycoprotein - useful for
PT recombinant protein production in high quantities.
PS Claim 1; Pages 16-18; 22pp; Japanese.
CC 090444 encodes R73993 hamster oviduct specific glycoprotein (HOGP).
CC The cDNA can be used for the commercial recombinant prodn. of
CC HOGP in high quantities.
SQ Sequence 2366 BP; 602 A; 574 C; 573 G; 617 T;

Query Match 4.28; Score 69; DB 15; Length 2366;
Best Local Similarity 61.18; Pred. No. 1.61e-27;
Matches 232; Conservative 0; Mismatches 145; Indels 3; Gaps 2;

Db 60 aagctgtgtcattccactgagctacagtcagtcgctccctccatctg 119

```

Oy 68 AACTGTCCTGCTACTTCACCAACTGGGCGGACAGACAGAGGGGAGGCGCTTCTCG 127  
Db 120 ccccgtagcagctgagatcccttcttctgtacacacatgattccttgcgcagatgagc 179  
Oy 128 CCCAAGGACTTGGACCCCGCTTTCACCCCGCTTATCTACGCCCTTCCTGGCAGAGAC 187  
Db 180 aacaatcagatgttctgcataatctccagatgagaataatctatccaatcagac 239  
Oy 188 AACCAACAGACTGAGCAGCCAGCTAG-TGA-ATGACGAGACTCTCTACAGAGATTCAAT 244  
Db 240 aaactcaagagagagagacagagccctgaaacactactctgttggagagctggaattc 299  
Oy 245 GGCCTGAAGAAATGATCAATCCCAAGCTGAGACCTGTTAGCCATCGAGCGCTGAATTTC 304  
Db 300 ggcacatcaggttccacacatgctgtccacccttgccagcgctgaaatatttgg 359  
Oy 305 GGCCTGAGAACTTTCAGATATGAGGACGAGCCGACACACCGTTCAGACTTTGTCAAC 364  
Db 360 teagttgtatctcctcctgagacacatggtcttggatggcttctcttctgtac 419  
Oy 365 TCGGCATCAGGTTCTCGCAATACAGCTTGAAGCGCTTGACCTTGAGCTGGAGATTC 424  
Db 420 cctgactacagagagagccc 439  
Oy 425 CCAGGAAGCCAGGGAGGCC 444

RESULT 8  
ID N81164 standard; DNA; 204 BP.  
AC N81164; 08-NOV-1990 (first entry)  
DE Base substituted E.coli beta-galactosidase alpha-fragment.  
KW E.coli beta galactosidase alpha-fragment; base substitutions; ss.  
OS Escherichia coli.  
FH Key location/Qualifiers  
FT misc-feature 19..69  
FT /\*tag- a  
FT /function-multiple cloning site  
FT primer\_bind 187..204  
FT /\*tag- b  
PN EP-285123-A.  
PD 05-MAY-1988.  
PF 30-MAR-1988; 105163.  
PR 03-APR-1987; US-034819.  
PI (SUSO) STUMEN SOKERI OY.  
PI Lehtovaara P, Knowles J, Kolivula A, Bamford J, Reinikainen T;  
DR WPI; 88-279927/40.  
PT Introducing random point mutations into nucleic acids -  
PT by prepn of single stranded template; annealing a primer, elongation,  
PT misincorporation, completion of molecules and screening.  
PS Disclosure; P: English.  
CC Random point mutations were introduced into the alpha fragment of  
CC E.coli beta-galactosidase. The wild type sequence was obtained as a  
CC single stranded template and an oligonucleotide was hybridised to  
CC it to generate a popn of DNA molecules which terminate at all  
CC possible nucleotide positions within a specified region. The  
CC variable 3' ends generated in this way are used as primers for  
CC reverse transcriptase. Nucleotides are misincorporated by the  
CC transcriptase and the molecules are completed to forms that can be  
CC amplified and then expressed in a suitable host-vector system.  
CC The sequence covers all 176 diff base substitutions, most of which  
CC occurred singularly in any given mutant.  
CC See also P80575.  
SQ Sequence 204 BP; 21 A; 47 C; 17 G; 11 T; 108 Others;

Query Match 3.1%; Score 51; DB 1; Length 204;  
Best Local Similarity 15.4%; Pred. No. 4.79e-16;  
Matches 18; Conservative 62; Mismatches 36; Indels 1; Gaps 1;

Db 75 yyedchvgcgytmrtthhyrmhvrvdyndmsdaayccyrrsvkydcocynachhd 134  
Oy 603 CCCAGAACTGTGATTTGCAACCTTATGCGCTACAGACTTCATGCTCTTGGAGAAAG 662

Db 135 hvvbybbyvnhnncncbnnhvcnvhbnhmvayrvhdariddvhcvcchc 191  
Oy 663 TCACGGACATATACAGCCCTC-TACAAGAGCAAGAAAGAGATGTGTGACAGACC 718

RESULT 9  
ID Q51746 standard; cDNA; 91 BP.  
AC Q51746;  
DT 31-MAY-1994 (first entry)  
DE Oligonucleotide probe MK14-A  
KW Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;  
KW ss.  
OS Synthetic.  
PN EP-571911-A.  
PD 01-DEC-1993.  
PR 24-MAY-1993; 108325.  
PR 26-MAY-1992; US-889651.  
PA (BECT ) BECTON DICKINSON CO.  
PI Shank DD, Spears PA;  
PI WPI; 93-378844/48.  
DR New oligo:nucleotide probes specific for Mycobacteria - used for  
PT detection and amplification of Mycobacteria nucleic acid in  
PT samples  
PS Claim 3; Page 14; 23pp; English.  
CC Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14  
CC (Q51735). It hybridized to all spp. of mycobacteria tested, but  
CC cross reacted to a few non-mycobacterial spp. The probe may  
CC be useful as an initial screen for mycobacterial infection.  
CC See also Q51735-45 and Q51747-59.  
SQ Sequence 91 BP; 5 A; 17 C; 15 G; 4 T;

Query Match 2.8%; Score 45; DB 9; Length 91;  
Best Local Similarity 2.0%; Pred. No. 2.21e-12;  
Matches 1; Conservative 47; Mismatches 3; Indels 0; Gaps 0;  
Db 10 gsvhsyyvvvnhshvshvnhvshvvvnhvshvvvnhvshvhyvsv 60  
Cp 1361 GGCTGTGGCGAGCTTGTGTGAACAGCCGCCCGCTGACAGCTGAGA 1311  
RESULT 10  
ID Q51746 standard; cDNA; 91 BP.  
AC Q51746;  
DT 31-MAY-1994 (first entry)  
DE Oligonucleotide probe MK14-A  
KW Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;  
KW ss.  
OS Synthetic.  
PN EP-571911-A.  
PD 01-DEC-1993.  
PR 24-MAY-1993; 108325.  
PR 26-MAY-1992; US-889651.  
PA (BECT ) BECTON DICKINSON CO.  
PI Shank DD, Spears PA;  
PI WPI; 93-378844/48.  
DR New oligo:nucleotide probes specific for Mycobacteria - used for  
PT detection and amplification of Mycobacteria nucleic acid in  
PT samples  
PS Claim 3; Page 14; 23pp; English.  
CC Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14  
CC (Q51735). It hybridized to all spp. of mycobacteria tested, but  
CC cross reacted to a few non-mycobacterial spp. The probe may  
CC be useful as an initial screen for mycobacterial infection.  
CC See also Q51735-45 and Q51747-59.  
SQ Sequence 91 BP; 5 A; 17 C; 15 G; 4 T;

Query Match 2.6%; Score 42; DB 9; Length 91;  
Best Local Similarity 3.8%; Pred. No. 1.37e-10;  
Matches 2; Conservative 45; Mismatches 5; Indels 0; Gaps 0;

Db 12 svhsyyvvvnhshvshvshvshvvvnhvshvvvnhvshvhyvsvctc 63  
Oy 124 CTTGCCAAGACTGTGAGCCCGCTTGTGACCCACCTCATCTACGCTTC 175







in, or flanking, the unpredicted or variant residues. These residues confer some degree of conformational rigidity to the peptides. The TSARs or compens. comprising a TSAR binding domain can be used in vivo to deliver a chemically or biologically active moiety, eg. metal ion, radioisotope, peptide, toxin or enzyme, to the specific target or on the cell. They can also replace the function of macromolecules, eg. monoclonal or polyclonal antibodies and therefore circumvent the need for complex methods of hybridoma formation or in vivo antibody production. The TSARs are easily characterised and have designed activity allowing direct and rapid detection in a screening process.

Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;  
Query Match 2.2%; Score 36; DB 12; Length 114;  
Best Local Similarity 3.6%; Pred. No. 4.09e-07;  
Matches 4; Conservative 34; Mismatches 74; Indels 0; Gaps 0;

Db 3 bnbdbnbdbnbdbnbtgcdbnbdbnbdbnbdbnbdbnbdbnnnnnn 62  
QY 130 CAAGGACTTGAGCCCGACGCTTGACCCGCTCATCTAGCCTCGCTGGCATGACCAA 189  
Db 63 bnbdbnbdbnbdbnbdbnbdbnbdbnbdbnbdbnbtgcdbnbdbnnb 114  
QY 190 CCACGAGCTGAGCACCACCTGAGTGAATGACGAGACTCTCTACGAGAGTTTC 241

Search completed: Sun Jun 29 19:58:47 1997  
Job time : 183 secs.





[illegible]

RESULT	2
ID	R73991 standard; Protein; 537 AA.
AC	R73991;
DT	10-JAN-1996 (first entry)
DE	Bovine oviduct specific glycoprotein.
KW	Bovine oviduct specific glycoprotein; recombinant production; BOGP.
OS	Bos taurus.
FX	Key
FT	Peptide
FT	/label= sig_peptide
FT	Peptide
FT	/label= mat_peptide
FN	007107979-A.
PD	25-APR-1995.
PF	15-AUG-1994; 214227.
PR	19-AUG-1993; JP-227881.
PA	(KINO-) KINOSEI PEPTIDE KENKYUSHO KK.
DR	WPI: 95-190179/25.
DR	N-PSDB: Q90442.
PT	New DNA encoding an oviduct-specific glycoprotein - useful for
PT	recombinant protein production in high quantities.
PS	Claim 4; Pages 7-9; 22pp; Japanese.
CC	O90442 encodes R73991 bovine oviduct specific glycoprotein (BOGP).
CC	The cDNA can be used for the commercial recombinant prodn. of
CC	BOGP in high quantities.
CC	Sequence 537 AA;

	Query Match	Similarity	50.4%	Score 1345	DB 145	Length 537
	Best Local	Matches 191	Conservative	Pred. No. 6.48e-115	Mismatches 101	Indels 7
						Gaps
Db	6	glllvkhhgdaahklyvcfwnafsrppgaspilprldcbflcthlvafafasmsnqivp	65			
Oy	9	gfmvlmifpwsaaklvocyfnnmaoqyogearflprkldslcthlilaragmttnqlstr	68			
Db	66	kdpqgklllypefnklykernrglktllslsgwmfgtvrftmstlsfserfvssviall	125			
Oy	69	twmndet-lyxfefnglkkmmnkeltlllaigcmngtorkfdymatannrogtfvnsairfl	127			
Db	126	rthgfdgdldfflypgljhsparprwfvflleelbqafkneagltmrrplllsaavsgd	185			
Oy	128	rkysidgldlwmepvsgsgspavdkerrtllvoolanafoeomtsekeklldlsaaavpag	187			
Db	186	phvvqkavearllgrldffisvlydldhgswekvctghnsplfs-1--pgdpbs-sa-ya	240			
Oy	188	qtyvdagcvgevkinkonndfvlmlaydfhgsmeaktgtnsplrkrqoesglaasltandaa	247			
Db	241	nywtqlygppekllmglygrtfnllkaasneelraqavpasbpkytkagafslayyeic	300			
Oy	248	qomlorkgpasklltigmvtygrsftllassdptrbgapatsggrpgpettekgmllayevc	307			
Db	301	cfvrrakkrvindygyyafkgykewygdaisfgyaafikrzhnfgamwtdldldfr	360			
Oy	308	sw-kgatqrordokvplifpndmowgdvdesekrtvysylkrokgloglaawmaldedfa	366			

```

Db      361  gyfcgtgpfplvhtlmnl 379
        | : | | : | : | | |
QY      367  GFSCNQGRYPPIQLRQEL 385

```

RESULT	3	
ID	R73992	standard; Protein; 718 AA.
AC	R73992.	
DT	10-JAN-1996	(first entry)
DE	Murine oviduct specific glycoprotein.	
KW	Murine oviduct specific glycoprotein; recombinant production; MOPF.	
OS	Mus musculus.	
FH	Key	Location/Qualifiers
FT	Peptide	1..18
FT	/label-sig_peptide	
FT	Peptide	19..718
FT	/label-mat_peptide	
PN	J07107979-A.	
PD	25-APR-1995.	
PF	15-AUG-1994; 214227.	
PR	19-AUG-1993; JP-227881.	
PA	(KINO-) KINOSEI PEPTIDE KENKYUSHO KK.	
DR	WPI: 95-180179/25.	
DR	N-PSDB: Q90443.	
PT	New DNA encoding an oviduct-specific glycoprotein - useful for recombinant protein production in high quantities.	
PS	Claim 4; Pages 11-14; 22pp: Japanese.	
CC	Q90443 encodes R73992 murine oviduct specific glycoprotein (MOPF).	
CC	The cDNA can be used for the commercial recombinant prodn. of MOPF in high quantities.	
SO	Sequence	718 AA;

	Query Match	38.7%	Score 1332:	DB 14:	Length 718:
	Best Local Similarity	49.2%	Pred. No.117e-113;		
	Matches 191; Conservative	81;	Mismatches 107;	Indels 9;	Gaps 6;
Dd	7 lvlhmshdcaaykrivcyftmwarhsrppasimphldapfictlhlifafasmnqiyak	66			
Oy	11 MVLMIIPW-GSAKILVCFTTMAQIRGEAKFLPRKDDLPSCITHLIYAFAGTNHQLSTT	69			
Dd	67 nlqdenlylpefnklkenre klllsigynfgtsrfamslstlaneckfdsvlsflr	126			
Oy	70 ENMDET-LYGEENGLKKNPKKLTLAIGNNFGRQKRTDVAATANNQOTVNSAIRLRL	128			
Dd	127 ihgfglalffilypglrgspdhrtwnffileelqfaferealltqbhrillsaavgip	186			
Oy	129 KYSFGPLDLMEWPESQSGPAVDKERFTTLLVDLANAFQOEKQTSGKRRLLSAAVPAPG	188			
Dd	167 slthsydaallgrrldfinvlsydlngswekfctghnsplfslpedsassa--y---am	241			
Oy	189 TYVDAGYEEDKLAQNLDPEVNLMAYDFHESWEKVGHNSPLKROEGSQAASLNDAVVQ	248			
Dd	242 ywrk grtpacklingfpvyrnfyllkesknlgiasmgpspygtkqaaflayevcs	301			
Oy	249 QWLQGTGFSKILLDEMPFYGRSEFTLIASSSDTRVGAPATGSCTPGPFTLEGGLATLEYCS	308			
Dd	302 fvgarkkhwidqvyppayafkgkwlgyddtisfsykanyvrrehngamwvtldmdvrg	361			
Oy	309 W-KGKTQRIDQDKPYPIFRDNQWGWGFVDVESFKTKVSYLQKGGLGAMNAALDIIDPDAG	367			
Dd	362 tfcgnqpfiplvnhl-nellvqtessnstp	388			
Oy	368 FSCNQGRRYPLIQTLRQELSLPYLPESGP	395			
RESULT	4				
ID	P81342 standard; protein; 321 AA.				
AC	P81342:				
DT	19-OCT-1990 (first entry)				
DE	Polypeptide involved in protective mechanisms				
KM	Immune response; cell growth.				
PN	J63032898-A.				

[illegible]

RESULT 5  
ID R70025 standard; Protein: 371 AA.  
AC R/0025;  
DT 29-SEP-1995 (first entry)  
DE Tobacco chitinase encoded by tobacco Cluster-A cDNA clone CA-3.  
KM Tobacco: chitinase; antifungal; fungicide; Cluster-A.  
OS Nicotiana tabacum.  
FH key Location/Qualifiers  
FT Peptide I..18  
FT /label= signal peptide  
FT Misc\_difference 22  
FT /note= "not present in genomic clone"  
FT Misc\_difference 182..195  
FT /note= "determined by protein sequencing"  
FT FT Misc\_difference 216...241  
FT /note= "Determined by protein sequencing"  
PN EP-639642-A.  
PD 22-FEB-1995.  
PF 17-AUG-1993; 202425.  
PR 17-AUG-1993; EP-202425.  
PA (MOGE-) MOGEN INT NV.  
PA (UYLE-) RIJKSWIV LEIDEN.  
PI Apotheker-de Groot M. Bol JF, Cornelissen BJC, Linthorst HJM,  
PI Melchers IS, Ponstein AS, Sela-bourlidge MB;  
DR WPI: 95-083454/12.  
DR DR  
PT New plant protein having endo-chitinase activity - used in  
PT antifungal compsns. and to develop transformed plants which are  
PT less susceptible to fungal infection.  
PS Claim 3: Page 19-20: 43pp. English.  
CC Screening of a lambda ZAP cDNA library of TMV-infected Samsumn NN  
CC tobacco plants with a probe derived from PRO840 (a partial Cluster-A  
CC cDNA clone) resulted in the isolated of 11 positively hybridising  
CC clones. Analysis revealed that all were identical and corresp. to  
CC Cluster-A cDNA. The nt sequence of clone CA-3 is given in 082973/  
CC R70025. This clone lacks 7 codons of the 5' part of the ORF. The  
CC cDNA clone was completed by performing a PCR reaction on clone CA-3  
CC with primer T7 (082794) and primer P1 (082975). A partial AA  
CC sequence of the protein isolated from TMV-infected tobacco leaves

CC	was obt'd. and found to be almost identical to the AA sequence	
CC	deduced from the cDNA.	
5Q	Sequence 371 AA;	
	Query Match 17.3%; Score 597; DB 13; length 371;	
	Best Local Similarity 33.0%; Pred. No. 2,35e-43;	
	Matches 107; Conservative 86; Mismatches 108; Indels 23; Gaps 16;	
Db	36 midsctltfnlfcadfnpsngllispengds-fsqfstvqrkpsvktlslagrr 94	
QY	44 KDLDSLTCHTIIYAFAGNT-NHQLSTEWDETLVDFEFGN-LKKMKPKTKTLAI-GGW 99	
Db	95 a-dttaygimargpsrkskfidsrlarqfghlqldewp-L-s-atdmtnlgll 149	
QY	100 NFGTKRFTDMWATANNKRTFVNSAIRFLRKTSFPGCLLDIDHEPSSQSPAVDKERFTTLV 159	
Db	150 newrtalmeaensgraaillltaavsyprvnglnypvesarlnlwnlmaydfipnw 209	
QY	160 QDLNAFAFOEQAQTSKEKELLTSAVPAGQFVVDAGYEVDKIAQLNLMAYDFHGS-W 218	
Db	210 spscqnsnaqlfdpvnh-lsgs--Ag-lhawiaqavppkklivlgfipfyawrlvnpn 263	
QY	219 EKV-TGHNSPIYKROESGAALSLNVDAAVOOMLQKSTPKSKLLGMPTRGSRSTLA 277	
Db	264 lndlrapaagksnvga-vddgsm-tynrirdyivqtatcvynatlyvgdycsqsmisay 321	
QY	278 DTRVCAPATGSGTGPFFRKKEGMLAYEVOSWKKATQKR-IDQQR-VP-YTFDNQWGF 334	
Db	322 ddtgsvrnkvnyvkgrygllyfaw 345	
QY	335 DVESEFKTVSYLKKQKGLGAMVW 358	

```

ID      R70029 standard; Protein: 377 AA.
AC      R70029;
DE      29-SEP-1995 (first entry)
DD      Tobacco cluster-A protein encoded by genomic clone.
KW      Tobacco; chitinase; antifungal; fungicide; Cluster-A.
OS      Nicotiana tabacum.
FH      Key
FT      Location/Qualifiers
FT      Peptide
FT      1..25
FT      /label= signal peptide
FT      Misc-difference 28..29
FT      /note="cDNA encodes additional ile here"
FT      Misc-difference 188..201
FT      /label= Determined by sequencing the protein
FT      Misc-difference 222..247
FT      /note="Determined by sequencing the protein"
FT      EP-639642-A.
PD      22-FEB-1995.
PN      17-AUG-1993; 202425.
PR      17-AUG-1993; EP-202425.
PA      (MOGE-) MOGEN INT NV.
PA      (UTLE-) RIJSDUNTJ LEIDEN.
PI      Apolteker-de Groot M, Bol JF, Cornelissen BJC, Lanthorst HJM;
PI      Melchers LS, Ponstein AS, Sela-buurlage MB;
PI      WPI: 95-083454/12.
DR      N-PSDB: 082976.
PT      New plant protein having endo-chitinase activity - used in
PT      antifungal compns. and to develop transformed plants which are
PT      less susceptible to fungal infection.
PS      Claim 3; Page 25-28; 43pp; English.
CC      Screening of a lambda xap cDNA library of TMV-infected Sansun NN
CC      tobacco plants with a probe derived from PRO84 (a partial Cluster-A
CC      cDNA clone) resulted in the isolated of 11 positively hybridising
CC      clones. Analysis revealed that all were identical and corresp. to
CC      Cluster-A cDNA. The sequence of cDNA clone ca-3 is given in 082973/
CC      R70029. A genomic library of N. tabacum was screened using the
CC      Cluster-A cDNA insert of clone ca-3 as a probe (see 082977, 082978).
CC      The complete nt. sequence of Cluster-A cDNA including the deduced
CC      primary structure of the Cluster-A protein the 5' and 3' UTR regions
CC      of the gene are shown in 082976/R70029. Comparison of the cDNA

```









[illegible]

RESULT 14  
ID R32547 standard; Protein: 238 AA.  
AC R32547;  
DT 17-JUN-1993 (first entry)  
DE HANPV IE-1 gene protein (partial sequence).  
KW Heliothis armigera nuclear polyhedrosis virus; Ha; NPV;  
KW Insecticide; immediate early-1.  
OS Heliothis armigera nuclear polyhedrosis virus.  
PN W09303144-A.  
PD 18-FEB-1993.  
PF 05-AUG-1992; AU0413.  
PR 05-AUG-1991; AU-007576.  
PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
PI Christian PD, Dali DJ, Gordon KHJ, Hanzlik TN, Srisankha A.  
DR WPI. 93-076801/09.  
DR N-PSDB: Q35882.  
PT Insect virus with reduced capacity to occlude viral particles - used  
PT for controlling proliferation of insect pests without horizontal  
PT transmission  
PS Disclosure: Fig 2; Sipp: English.  
CC The sequence is that encoded by the Heliothis armigera (Ha)  
CC nuclear polyhedrosis virus (NPV) immediate early-1 (IE-1) gene  
CC (partial sequence).  
SQ Sequence 238 AA;

Query Match	7.9%;	Score 271;	DB 6;	Length 238;
Best Local Similarity	34.1%;	Pred. No. 1.65e-13;		
Matches	58;	Conservative	43;	Mismatches 51;
				Indels 18;
				Gaps 14

D<sub>b</sub> 44 ercdfflanpygrveflwtkf-fdgldldwefpgkgaanpnvgdverdnmtlytlajlela 102  
 Q<sub>y</sub> 115 NRQFVNSAI-R-F-L-RKSFSDLDDWEPGSGS-PAY-DKER-FTTLVODLAN 164  
 D<sub>b</sub> 103 mldg-vqdgtrrtleltalsagldktaav-ncdraagqyldkifmsydkdkgawntldg 160  
 Q<sub>y</sub> 165 AFQOEAQTSGRERLLLSAAYPAGQGYVDAGYEVDKIAQNIDFVNLMAYDFHSGMERKV-G-223  
 D<sub>b</sub> 161 hgtaly--gsaknpnepytanvaadalaaqyvp-kklvgvamygwt 207  
 Q<sub>y</sub> 224 HNSPLRKROESGAASLNDAAVQOWL-RKGPRASLLIGMPTGYSFT 272

RESULT	ID	15
AC	W02156	standard; Protein; 1046 AA.
AC	W02156;	
DT	14-JAN-1997	(first entry)
DE	Periplasmic chitodextrinase.	
KW	Periplasmic chitodextrinase; periplasmic Beta-N-acetylglucosaminidase;	
KW	Beta-N-acetylglucosaminidase; chitin, oligosaccharide; catabolic;	
KW	catabolism.	
OS	Vibrio furnissii.	
PN	W06925424-AL.	
PD	22-AUG-1996.	
PF	13-FEB-1996; U023332.	
PR	13-FEB-1995; US-386727.	
PA	(UOYO ) UNIV JOHNS HOPKINS.	
PI	Baszler B, Chittlaru E, Keyhani N, Roseman S, Rowe C;	
PI	Yu C;	
DR	WPI; 96-393335/39.	
DR	N-PSDB; T36387.	
DR	Chitin blosynthetic enzymes end I, exo I and exo II - are	

periplasmic chito:dextrinase(s), periplasmic beta-glucanidase(s) and Pr ariI beta-N-acetyl:glucosaminidase(s), respectively  
PS Claim 2; Page 68-71; 101pp; English.  
CC Periplasmic chitodextrinase (W02156), periplasmic  
CC Beta-N-acetyl:glucosaminidase (W02157) and ariI  
CC Beta-N-acetyl:glucosaminidase (W02158) can be used to produce chitin  
CC oligosaccharides with the structure (GlcNAc)<sub>n</sub> where n is 2 or  
CC higher, by contacting them with soluble chitin. The enzymes are  
CC encoded by the genes endl, exoi and exoii respectively. They are  
CC all genes involved in the catabolic pathway of chitin.  
SQ Sequence 1046 AA;

Query Match	6.48;	Score 219;	DB 19;	Length 1046;
Best Local Similarity	33.38;	Pred. No. 5.16e-09;		
Matches	48;	Conservative	40;	Mismatches 42;
			Indels 14;	Gaps 10;

Dd		454 etfadsavemmrxkyrfqldidrltsnldgctgnppddatfeesrraylmsyheilmvr	513
OY	117	QTFVNSARFRPKRYSFDDLD-D-W--EYPGSGSPD-PA--KE-RFTTLY---OPLANAFQ	167
Dd		514 eklvaasagdvhymltaaps-ayllrmetnavcgyldvrylmasydlhgvanhvqh	574
OY	168	GDAQ-TSQR-K--RLLSAIVPAGQTYYDAGEVDKIAQNDFVNIMAYDEFGSWEKVTGH	224
Dd		573 naalydtgkdselagw-nvytqaq	595
OY	225	NSPLYKRQEESGAALSNVDAAQ	248

Search completed: Fri Jun 27 16:54:04 1997  
Job time : 116 secs.



\*\*\*\*\*  
Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.  
\*\*\*\*\*  
MARCH protein - protein database search, using Smith-Waterman algorithm  
Run on: Fri Jun 27 16:49:12 1997; Maspar time 21.48 seconds  
Tabular output not generated. 618.601 Million cell updates/sec  
Title: >US-08-663-618A-2  
Description: (1-466) from US08663618A.pep  
Perfect Score: 3443  
Sequence: 1 MVRSVAMAGFVNLMPWGS.....QQSPTGLVFNSSCKCCTWN 466  
Scoring table: PAM 150  
Gap 11  
Searched: 89912 segs, 28507787 residues  
Post-processing: Minimum Match 0%  
Listing first 45 summaries  
Database: plr50  
1:ann1 2:ann2 3:ann3 4:ann4 5:annan1 6:annan2 7:annan3  
8:annan4 9:annan5 10:annan6 11:annan7 12:annan8  
13:annan9 14:annan10 15:unrec 16:unrev  
Statistics: Mean 48.408; Variance 102.872; scale 0.471

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	1519	44.1	383	13	A49562 cartilage glycoprotein	8.84e-265
2	1462	42.5	383	16	heparin-binding glyco	1.71e-253
3	1459	42.4	381	14	BRP39 protein - mous	6.73e-253
4	1459	42.4	381	14	BRP39 protein - mous	6.73e-253
5	1380	40.1	399	5	secretory protein YM	2.80e-237
6	1361	39.5	539	14	estrogen dependent o	1.58e-233
7	1345	39.1	537	14	oviduct-specific gly	2.27e-230
8	1185	34.4	483	12	chitinase (EC 3.2.1.	7.24e-179
9	1086	31.5	504	12	chitinase (EC 3.2.1.	1.85e-179
10	1031	29.9	554	12	chitinase (EC 3.2.1.	1.04e-168
11	698	20.3	699	10	chitinase (EC 3.2.1.	2.46e-104
12	594	17.3	378	5	chitinase (EC 3.2.1.	1.33e-84
13	582	16.9	424	5	chitinase (EC 3.2.1.	2.42e-82
14	576	16.7	423	5	chitinase (EC 3.2.1.	3.26e-81
15	561	16.3	427	12	chitinase (EC 3.2.1.	2.14e-78
16	546	15.9	427	12	chitinase (EC 3.2.1.	1.38e-75
17	446	13.0	563	16	chitinase precursor	4.70e-57
18	404	11.7	561	9	chitinase (EC 3.2.1.	2.11e-49
19	344	10.0	799	8	chitinase (EC 3.2.1.	1.19e-38
20	345	10.0	799	8	chitinase (EC 3.2.1.	7.93e-39
21	344	10.0	820	16	chitinase (EC 3.2.1.	1.19e-38

22	337	9.8	499	9	S04856	chitinase (EC 3.2.1.	2.06e-37
23	314	9.1	511	11	S61156	probable membrane pr	2.24e-33
24	287	8.3	610	10	JH0573	chitinase (EC 3.2.1.	1.06e-28
25	266	7.7	452	12	JC4038	47K glycoprotein - f	4.02e-25
26	263	7.6	1146	11	S07915	R22 protein - yeast	1.29e-24
27	246	7.1	413	5	JC2135	chitinase (EC 3.2.1.	9.11e-22
28	229	6.7	124	5	S57715	chitinase (EC 3.2.1.	5.81e-19
29	172	5.0	36	14	A27682	33K whey protein - b	5.52e-10
30	172	5.0	83	13	A57954	estradiol-stimulated	5.52e-10
31	136	4.0	29	14	S57204	oviduct-specific gly	8.32e-05
32	136	4.0	29	14	S55693	oviduct-specific sta	8.32e-05
33	121	3.5	597	5	S33848	chitinase (EC 3.2.1.	8.07e-03
34	121	3.5	597	5	S32039	chitinase (EC 3.2.1.	8.07e-03
35	112	3.3	642	16	S37623	homologues with orf	1.08e-01
36	110	3.2	488	10	A41961	chitinase (EC 3.2.1.	1.90e-01
37	104	3.0	305	9	S28440	probable coproporphy	9.80e-01
38	104	3.0	422	4	ZWEC	unuc protein - Esche	9.80e-01
39	102	3.0	662	11	S61193	probable membrane pr	1.67e+00
40	105	3.0	787	8	PC1232	retrovirus-related c	7.49e-01
41	105	3.0	1409	4	OFECF	retrovirus-related p	7.49e-01
42	101	2.9	562	3	KRHDB	keratin, 56K type II	2.17e+00
43	101	2.9	958	1	JN0102	glucan 1,4-alpha-glu	2.17e+00
44	101	2.9	1045	2	SUSEM	serine proteinase (E	2.17e+00
45	101	2.9	1070	11	S19686	alpha-glucosidase (E	2.17e+00

## ALIGNMENTS

RESULT 1  
ENTRY 1  
TITLE A49562  
ALTERNATE\_NAMES cartilage glycoprotein gp33 precursor - human  
ORGANISM 39K synovial protein  
#formal\_name Homo sapiens #common\_name man  
DATE 23-Mar-1995 #sequence\_revision 23-Mar-1995 #text\_change 03-May-1996

ACCESSIONS  
REFERENCE A49562; S10677; A33162  
#authors Hakala, B.E.; White, C.; Recklies, A.D.  
#journal J. Biol. Chem. (1993) 268:25803-25810  
#title Human cartilage gp-39, a major secretory product of articular chondrocytes and synovial cells, is a mammalian member of a chitinase protein family.

#accession A49562  
#status preliminary  
#molecule\_type mRNA  
#residues 1-383 #label HAK  
#cross-references GB:M80927  
REFERENCE S10677  
#authors Nylkios, P.; Golds, E.E.  
#journal Biochem. J. (1990) 269:265-268  
#title Human synovial cells secrete a 39 kDa protein similar to a bovine mammary protein expressed during the non-lactating period.

#cross-references M80927; S10677  
#accession S10677  
#molecule\_type protein  
#residues 22-40, 'X', 42-45 #label NY2  
KEYWORDS cartilage; extracellular protein; glycoprotein  
FEATURE  
1-21 #domain signal sequence #status predicted #label SIG  
22-383 #product cartilage glycoprotein gp39 #status predicted  
#label MAT

SUMMARY #length 383 #molecular-weight 42613 #checksum 2942

Query Match 44.1%; Score 1519; DB 13; Length 383;  
Best Local Similarity 52.8%; Pred. No. 8.84e-265;  
Matches 201; Conservative 88; Mismatches 83; Indels 9; Gaps 5;

Db 8 tgfvtvllgcsayklycytswsgregdscfpalgrfcthtlysfansndhd 67  
:|||||: |||||:|||||:| || |||||:|||||:| :| :|  
Oy 8 AGFVFLMIPWGSAAKIVCFYTMAGYRGEARLPKDDPSCTHTLYAFAGTNNQLS 67  
Db 68 twewndvtlygmintknrpnltllsvgwmfgsgqrfisklansltqsrfftkvppfl 127

QY	68	TTEWNNDELTYOEFNGLKKMNKLTLLAIGWNGCTOKFPDWMATANNRTPVNSAIRREL	127
Db	128	rlthfgdglawlypgrr-----dkqhftllkemkaefikeag-pgkqgllsaalsag	181
QY	128	RKYSFDGIDLMEWEPGSGSPAVDKERFTLLVODLANAFQOEAOTSGKERLLLSAAWPAG	187
Db	182	kvltdsydaksigshldffslmcyddfhgawrgtltphsrlpfsgqedaasdrtsncdyav	241
QY	188	QTYVDAGEVDKIAQNIDFNVLNMYDRHSGWEKTKVSNPLRYKROESGGAASLNDPAAY	247
Db	242	gymrlgapaaklmgjptfgtrfclass-etvgapaispgjipgrtkeagtlayelc	300
QY	248	QOMLOKGRPAKLLIGMPTYGRSTLTASSSDTRKGAATSGTPGPTKGGMLATYEV	307
Db	301	dfligatvhtlrgqvyatkngwvgyddqesvkskvyjlkdrqlaglanvwaljlddftg	360
QY	308	SW-KGATKQRIRIQDQKVPYIFRDNQWVGPDVESHKTKVSYLKKQGGAGAWMALDDDF	366
Db	361	gsfgqgdrlrflptalkdala	361
QY	367	GFSCNQG-RYPLIQTLRQELS	366
RESULT	2		
ENTRY	5S1327	#type complete	
TITLE	heparin-binding glycoprotein 38k - pig		
ORGANISM	#format_name Sus scrofa domestica	#common_name domestic pig	
DATE	07-May-1995	#sequence_revision 07-May-1995	#text_change 07-May-1995
ACCESSIONS	5S1327		
REFERENCE	5S1327		
Authors	Shackelton, L.M.; Mann, D.M.; Mills, A.J.T.		
#authors	3		
#description	Submitted to the EMBL Data Library, January 1995		
#description	Identification of a 38kDa heparin-binding glycoprotein (gp38k) in differentiating vascular smooth muscle cells as a member of a group of proteins associated with tissue remodeling.		
#accession	5S1327		
#status	preliminary		
#residues	1-383	#label SHA	
#cros-references	EMBL:247803		
SUMMARY	#length 383	#molecular-weight 42443	#checksum 2907
Query Match	42.5%;	Score 1462;	DB 16;
Best Local Similarity	51.0%;	Pred. No. 1.71e-253;	
Matches	198;	Conservative 87;	Mismatches 94;
		Indels 9;	Gaps 5;
Db	1	mglyvaagqgfayalvlgscaykilycytswsqyrgedgscfpdaipfclthlysfan	60
QY	1	WVRSYAMAGFVNLIMPGSAAKLVICYFTTMAQYRQGEARLPKRDPSICLHLIYAFAG	60
Db	61	lsmaeidlwmvdlvlydntlknrpnlkltlsvgwmgfsgqfksiasntgsrrftl	120
QY	61	MTNHLSTTEWNNDELTYOEFNGLKKMNKLTLLAIGWNGCTOKFPDWMATANNRTPV	120
Db	121	ksvpfllrthfgdglawlypgrr-----dkrhlrtlkemkaefyreal-pyterlll	174
QY	121	NSAIRFLKRYSPDGLDMEWEPGSGSPAVDKERFTLLVODLANAFQOEAOTSGKERLL	180
Db	175	sgvasagkvaidrgydiaqslqnlldfsllydfdhgawrgtltghsrlpfsgqdaasdrf	234
QY	181	SAAVPAGOTYDAGEVDKIAQNIDFNVLNMYDRHSGWEKTKVSNPLRYKROESGGAAS	240
Db	235	snadgavsyvrlrgapanklmgjptfgtrfclassk-tdivgaaspgjipgrtkeagtl	293
QY	241	LNVDAVQOMLOKGRPAKLLIGMPTYGRSTLTASSSDTRKGAATSGTPGPTKGGM	300
Db	294	layelcldlfgatvhtlrgqvyatkngwvgyddqesvksknykakyllksrqdalanvwt	353
QY	301	LAYEVCSW-KGATKQRIRIQDQKVPYIFRDNQWVGPDVESHKTKVSYLKKQGGAGAWMA	359
Db	354	ldlddfrgnfcognrlrflptalkdavl	361

QY	360	LDLDDFAFGSCNG-RYPLQITRLRQELS	386
RESULT	3		
ENTRY	S61550	#type complete	
TITLE	BRP39 protein - mouse		
ORGANISM	#formal_name Mus musculus	#common_name house mouse	
DATE	10-Apr-1996	#sequence_revision 19-Apr-1996	#text_change 03-May-1996
ACCESSIONS	S61550		
REFERENCE	S61550		
#authors	Morrison, B.W.		
#submission	submitted to the EMBL Data Library, November 1995		
#accession	S61550		
#status	preliminary		
#molecule_type	mRNA		
##residues	1-381	##label MOR	
##cross_references	EMBL:X93035		
SUMMARY	#length 381	#molecular_weight 43001	#checksum 5107
Query Match	42.4%	Score 1459;	DB 14; Length 381;
Best Local Similarity	49.9%	Pred. No.6,73e-253;	
Matches	194;	Conservative 88;	Mismatches 96; Indels 11; Gaps 9;
Db	1	mgmraaltgavimljqscaayklycyfswsgyevgsfipdaigpflicthiysfan	60
QY	1	MYSRVANAGFMVILMIPMGSAKLVICYFTNMAQYRGAEARFLPKDLPSLCTHLIYAFAG	60
Db	61	issdmstewmdesnydklnlkrntnltltsvvgwkfgckrrfseisntertaf	120
QY	61	MT-NHQSTSEWDELTLYOEFNGLKMKANRLKTLTLLAIGWNGTQKFTDVAATNNROTf	119
Db	121	vrsvafliysygfagldavlp-r-f-r-dkyfsctlikelnaeftkevg-pgrekl	174
QY	120	VNSAIRLRKRSFDGDLDMWEYSGSGSAVDKERTFTIYODLAAAFQEQTSCKERLL	179
Db	175	lsaaalsgkvaiddgydiaglaqhlidflnltmydfthgvrqtlgthspdlfggqkdtfr	234
QY	180	LSAALVPAGQTYVDAGYEVDRKIAONLDFVIMLMAVDFHGSWEKVTGHNSPLYKROESCAA	239
Db	235	ysrnyrvygmrligagaskllnglittfsgkftlass-englqaplsesglpygftkeag	293
QY	240	SLNDAVAVQWQLQGIPTASRLILGMPTYRSFTLASSSDTRVGADATSGTPGPTREGG	299
Db	294	tlayvelcdflkgeavhrlnenkyepfatkgynwgyvghesvknkvgflkcklagamwv	353
QY	300	MLATYEVCSW-KGATKTRIODQVPIPIFDNDQWGFDDVESKTKTVSYLKKQGLAGAMW	358
Db	354	aldlddfgg-tcqpkefiftlnaikdala	381
QY	359	ALDLDFAFGSCNGGRY-PLIOTRLRQELS	386
RESULT	4		
ENTRY	I48271	#type complete	
TITLE	BRP39 protein - mouse		
ORGANISM	#formal_name Mus musculus	#common_name house mouse	
DATE	02-Jul-1996	#sequence_revision 02-Jul-1996	#text_change 02-Jul-1996
ACCESSIONS	I48271		
REFERENCE	I48271		
#authors	Morrison, B.W.; Leder, P.		
#journal	Oncogene (1994) 9:3411-3426		
#title	neu and ras initiate murine mammary tumors that share genetic markers generally absent in c-myc and int-2-initiated tumors.		
#cross_references	MUTID:95060797		
#accession	I48271		
#status	preliminary; translated from GB/EMBL/DBJ		
#molecule_type	mRNA		
##residues	1-381	##label RES	
##cross_references	EMBL:X93035; NID:g1085065; CDS_PTD:g1085066		

GENETICS  
#note  
SUMMARY

gene name brp39  
#length 381 #molecular-weight 43001 #checksum 5107

Query Match 42.4%; Score 1459; DB 14; Length 381;  
Best Local Similarity 49.9%; Pred. No. 6,73e-253;  
Matches 194; Conservative 88; Mismatches 96; Indels 11; Gaps 9;

Db 1 mgmraalfavlmllqscgsklyvcsqyregvgsflpdaipfictllysfan 60  
1 mrvsvamafmvlmmpwmsaaklvcyftmwaqyrgoeafpdkldpdlcchllyafag 60  
Qy 1 mrvsvamafmvlmmpwmsaaklvcyftmwaqyrgoeafpdkldpdlcchllyafag 60

Db 61 isadmlscwewndesnydklnktrtnlklvlsgvqkfgekrfseasnterfat 120  
61 mt-nholsttemwderlyqefnlgkkmprkltilaigwnfctokftdmvatanrqr 119  
Qy 61 mt-nholsttemwderlyqefnlgkkmprkltilaigwnfctokftdmvatanrqr 119

Db 121 vsravapflsygfdglawlyp-r-lr-dkyfscllkelnaefkevq-pgrckll 174  
120 vnsaifrlkrkysfhdglmewyrgsgspavdkerfttlviodlanfogeagtsgerell 179  
Qy 120 vnsaifrlkrkysfhdglmewyrgsgspavdkerfttlviodlanfogeagtsgerell 179

Db 175 lsaalsagkvaldtygdiaqiaqhlidfinlmydfhgvwrqitqhsplfgqkdrfdr 234  
180 lsaavpagoityvdagyevdkiqonldfvnlmavdfhgswekvtghnsplykrqeesgaaa 239  
Qy 180 lsaavpagoityvdagyevdkiqonldfvnlmavdfhgswekvtghnsplykrqeesgaaa 239

Db 235 ysnvnyavgymlrlgaqaskllmqipfctgskftlaas-englqaplsgeglpgrfkeag 293  
240 slnvdaavoomlqoktrpaskllgmptyrgrsftllasssdrvgaprtsgtppgpfkeeg 299  
Qy 240 slnvdaavoomlqoktrpaskllgmptyrgrsftllasssdrvgaprtsgtppgpfkeeg 299

Db 294 clayaelcdfllygaevhrlsnekvfakqngvgyvghesvknkvglflekklagaww 353  
300 mlayevcsw-k-gatkoqiodkvpyfrdnqmwgfddvesfkvtksylkqkglagaww 358  
Qy 300 mlayevcsw-k-gatkoqiodkvpyfrdnqmwgfddvesfkvtksylkqkglagaww 358

Db 354 alldldfgr-tcqpkefflptnaikdala 381  
359 alldldfgr-tcqpkefflptnaikdala 386  
Qy 359 alldldfgr-tcqpkefflptnaikdala 386

RESULT 5  
ENTRY S27879 #type complete  
TITLE secretory protein YM-1 precursor - mouse  
ORGANISM Mus musculus #common\_name house mouse  
DATE 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 03-May-1996

ACCESSIONS S27879  
REFERENCE S27879  
#authors Chang, N.C.A.; Liu, C.H.; Chang, A.C.  
#submitters submitted to the EMBL Data Library, June 1992  
#description Molecular characterization of a secretory protein (YM-1) transiently expressed by activated murine peritoneal macrophages.

#accession S27879  
#molecule\_type mRNA  
#residues 1-399 #label CHA  
#cross-references EMBL:M94584

CLASSIFICATION #superfamily Streptomyces chitinase chl40  
FEATURE 1-21  
22-399 #domain signal sequence #status predicted #label SIG  
#product secretory protein YM-1 #status predicted #label MAY

SUMMARY #length 399 #molecular-weight 44528 #checksum 3208

Query Match 40.1%; Score 1380; DB 5; Length 399;  
Best Local Similarity 46.4%; Pred. No. 2.80e-237;  
Matches 181; Conservative 94; Mismatches 112; Indels 3; Gaps 3;

Db 121 lgvirflrqnfdglnldaqygsrsgpdkhlfsvlvkemrkafeesvekdiprll 180  
120 vnsaifrlkrkysfhdglmewyrgsgspavdkerfttlviodlanfogeagtsgerell 179  
Qy 120 vnsaifrlkrkysfhdglmewyrgsgspavdkerfttlviodlanfogeagtsgerell 179

Db 181 lscgagllidvksgrtsrncslldyigvmytdlndpdkdytgensplykspdydiksa 240  
180 lsaavpagoityvdagyevdkiqonldfvnlmavdfhgswekvtghnsplykrqeesgaaa 239  
Qy 180 lsaavpagoityvdagyevdkiqonldfvnlmavdfhgswekvtghnsplykrqeesgaaa 239

Db 241 dlnvdsllsywkdhaaseklvfgfpaqhtllsdpsktvgisaptlstcpqpktydeeg 300  
240 slnvdaavoomlqoktrpaskllgmptyrgrsftllasssdrvgaprtsgtppgpfkeeg 299  
Qy 240 slnvdaavoomlqoktrpaskllgmptyrgrsftllasssdrvgaprtsgtppgpfkeeg 299

Db 301 llayevcflnegatewvdpagvayagnewgyddvrsfklkagylkdmnlgavv 360  
300 mlayevcsw-k-gatkoqiodkvpyfrdnqmwgfddvesfkvtksylkqkglagaww 357  
Qy 300 mlayevcsw-k-gatkoqiodkvpyfrdnqmwgfddvesfkvtksylkqkglagaww 357

Db 361 wplmddfsgsfchgrhfrpftstlsgdlnl 390  
358 waldddfsfhgrhfrpftstlsgdlnl 387  
Qy 358 waldddfsfhgrhfrpftstlsgdlnl 387

RESULT 6  
ENTRY I46470 #type complete  
TITLE estrogen dependent oviduct protein precursor - sheep  
ORGANISM Ovis montanus #common\_name domestic sheep  
DATE 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 16-Aug-1996

ACCESSIONS I46470  
REFERENCE I46470  
#authors Desouza, M.M.; Murray, M.K.  
#submitters Endocrinology (1995) 136:2485-2496  
#title An estrogen-dependent secretory protein, which shares identity with chitinases, is expressed in a temporally and regionally specific manner in the sheep oviduct at the time of fertilization and embryo development.

#cross-references MUID:95269691  
#accession I46470  
#status preliminary; translated from GB/EMBL/DBD  
#molecule\_type mRNA  
#residues 1-539 #label DES  
#cross-references EMBL:U16719; NID:G985600; CDS:PID:G985601

SUMMARY #length 539 #molecular-weight 59555 #checksum 2120

Query Match 39.5%; Score 1361; DB 14; Length 539;  
Best Local Similarity 48.8%; Pred. No. 1.58e-233;  
Matches 189; Conservative 85; Mismatches 106; Indels 7; Gaps 5;

Db 1 mgklllvglmlmklhghaahklycyfenvafsrpgasllprldpflchlvafas 60  
1 mrvsvamafmvlmmpwmsaaklvcyftmwaqyrgoeafpdkldpdlcchllyafag 60  
Qy 1 mrvsvamafmvlmmpwmsaaklvcyftmwaqyrgoeafpdkldpdlcchllyafag 60

Db 61 mnnngivpdkpidekllypfnklkernrglktllsvgwnfgtsrftmstfnref 120  
61 mnnngivpdkpidekllypfnklkernrglktllsvgwnfgtsrftmstfnref 120  
Qy 61 mnnngivpdkpidekllypfnklkernrglktllsvgwnfgtsrftmstfnref 120

Db 121 vksviallrlthgfdglidfllypqlrsgpdrwtfvllleelqafkneagltmrrll 180  
120 vnsaifrlkrkysfhdglmewyrgsgspavdkerfttlviodlanfogeagtsgerell 179  
Qy 120 vnsaifrlkrkysfhdglmewyrgsgspavdkerfttlviodlanfogeagtsgerell 179

Db 181 lsaavsgdphvqkaydalllgrlldfslvslldhgswekvtghnsplykspdydiksa 240  
180 lsaavpagoityvdagyevdkiqonldfvnlmavdfhgswekvtghnsplykrqeesgaaa 239  
Qy 180 lsaavpagoityvdagyevdkiqonldfvnlmavdfhgswekvtghnsplykrqeesgaaa 239

Db 241 -yam--sy--wrglgyvpeklmqiplygrfthllrasqnelagagaagpkytykqg 295  
240 slnvdaavoomlqoktrpaskllgmptyrgrsftllasssdrvgaprtsgtppgpfkeeg 299  
Qy 240 slnvdaavoomlqoktrpaskllgmptyrgrsftllasssdrvgaprtsgtppgpfkeeg 299

Db 296 flayevcsfvgrakrwnldgvyvafkgykewvyddalsfygafkrlkrehfggaww 355  
300 mlayevcsw-k-gatkoqiodkvpyfrdnqmwgfddvesfkvtksylkqkglagaww 358  
Qy 300 mlayevcsw-k-gatkoqiodkvpyfrdnqmwgfddvesfkvtksylkqkglagaww 358

Db 356 tldldfgrnfcgtpfplahlnl 382

OY	359	ADDLDDFAFGSCNGRYPILQTLROEL	385
RESULT	7	S57197	#type fragment
ENTRY		oviduct-specific glycoprotein	95k precursor - bovine
TITLE		(fragment)	
ORGANISM		#formal_name Bos primigenius taurus	#common_name cattle
DATE		28-Oct-1995	#sequence_revision 03-Nov-1995 #text_change 03-May-1996
ACCESSIONS		S57197	
REFERENCE		S57197	
#authors		Sendai, Y.; Abe, H.; Kikuchi, M.; Satoh, T.; Hoshi, H.	
#journal		Biol. Reprod. (1994) 50:927-934	
#file		Purification and molecular cloning of bovine oviduct-specific glycoprotein.	
#accession		S57197	
##status		preliminary	
##molecule_type		mRNA	
##residues		1-537	##label SEN
##cross-references		EMBL:D16639	
SUMMARY		#length 537	#checksum 1695
Query Match		39.1%	Score 1345; DB 14; Length 537;
Best Local Similarity		50.4%	Pred. No.2.27e-230;
Matches	191;	Conservative	80; Mismatches 101; Indels 7; Gaps 6;
Db	6	gllvlvkhhdgaahklvcyftnwafsrpaspallprldpflclthlyfafasmsnglvp	65
OY	9	GFWVLIMPMGSAKLVCFFTNMAOYRGSEARFLPNLDPSLCIHLIYAFAGMTNHQLST	68
Db	66	kdpqdekillypefnklkerngylkllslsgwvfgrvtftmlstfsnerfssvlaill	125
OY	69	TEWMDET-LYOENGGKKMKPKLTLLALIGGWNEGTCQKFDMVATANNROTFFVSARFL	127
Db	126	rthgfagdalfflypgllgsparrdwtcfvlllelllqafneagltmrplllsaavgd	185
OY	128	KRTSPDGLDWDFYFESSQSPPAVDKERFTTLLVDLANAFOEQMOTSGKERLLLSAAPAG	187
Db	186	phvvqkayeearllgrlldtsvisydlhswekvtnghnspfls--pgdps-sa_yam	240
OY	188	GYVVDAGYEVDKLAQNLDLPFNLMAYDPHGSMENVTGINSPLIKRQESGAASLNDVAIV	247
Db	241	nywrqlgvppeklmg1ptygrtfnhlkasquelraqavgpaspkytkqagfiayeyic	300
OY	248	OOMLAKGTFASTKLITLMPTRYGRSFITLASSSDTRVGAPATSGPFPPTKGMLAYEVC	307
Db	301	cfrvrkkrrvwindqvpyafkgykwgyddaisfykaftikrnfhgannwcllddfir	360
OY	308	SW-KGATKRIIDOKPYIFRDNQWGFDVEEFKTKVSLKGLAGAMWALDLDDFA	366
Db	361	gyfcgctgpfplvntlnll	379
OY	367	GFSCNGRRYPILQTLROEL	385
RESULT	8	A53918	#type complete
ENTRY		chitinase (EC 3.2.1.14)	precursor - braconid wasp (Chelonus sp.)
TITLE		chitinase (EC 3.2.1.14)	
ORGANISM		#formal_name Chelonus sp.	
DATE		28-Jul-1995	#sequence_revision 28-Jul-1995 #text_change 03-May-1996
ACCESSIONS		A53918	
REFERENCE		A53918	
#authors		Krishnan, A.; Nair, P.N.; Jones, D.	
#journal		J. Biol. Chem. (1994) 269:20971-20976	
#file		Isolation, cloning, and characterization of new chitinase stored in active form in chitin-lined venom reservoir.	
#accession		A53918	
##status		preliminary	
##molecule_type		mRNA	

##residues	1-483	##label	KRI
##cross-references	GB:U10422		
KEYWORDS	glycosidase; hydrolase		
SUMMARY	#length 483 #molecular-weight 52013 #checksum 5938		
Query Match	34.4%;	Score 1185;	DB 12; Length 483;
Best Local Similarity	42.3%;	Pred. No. 7,24e-199;	
Matches	159; Conservative	100; Mismatches 100;	Indels 17; Gaps 11;
Db	19	aspmhvcyfigawsvyrgngkfidingldptlchlliysfvyngkdvkylpdsdipgn	78
Qy	19	GSAAKLVCYFFNNMAQYRGGEARFLPKDLDSLCHLLIYAFAGMTNHOLSTTE-WND-ET	75
Db	79	ldgfgfslrkkhpsvklmawvgmagsvpsfgmasdgateafagvnykflgqyqfd	138
Qy	76	L--YQEFGLKKMPPKLTLLAIGMFGTQKFTDMVATNNRQTFVNSAIRLKRYSFD	133
Db	139	gfddwepaqrgspadvkmvklckalkkafvq--h---d-yllsaavaeptsask	191
Qy	134	GLIDDMEXPSGSGSPAVDKERFTTLVGLDNLNNAFQGEAQSGKERLLLSAIVPAGQRYVDA	193
Db	192	sydiiaemsgyldfinlmtydthgpdwhtcmhappssahdsngelnklnvaavkylvlqn	251
Qy	194	GEYDKRAQNDLFFVNLAWAYDFHSGWEMVETGHNLSLYKROESEAALINDAAVVOOMLOK	253
Db	252	gvpkexllvvgpaygkfstlenspknkglgavpsagatgagdytgenllgyneiacemqkag	311
Qy	254	GTPASKLLKSGPTYGRSTTLAASSDTRVGAAPATSGTPPGFTREGGLATYEVCSW-KGA	312
Db	312	dwevvdnehkvpvaykxngvwsfddlaalakakagfikqeglysgamwsietdfkgl-c	370
Qy	313	TKQRIOD-QK-VPRIFRDMQWVGSDVDESPKTSKYSLYKQKGLGAGWALDLDDPAFGSC	370
Db	371	ge-kyvpalkalnsvlg	385
Qy	371	NGGRPLTLQTLRQELIS	386
RESULT	9		
ENTRY	A38221	#type complete	
TITLE	chitinase (EC 3.2.1.14) MFI - nematode (Brugia malayi)		
ORGANISM	Formal_name Brugia malayi		
DATE	04-Mar-1993 #sequence-revision 18-Nov-1994 #text-change		
ACCESSIONS	A38221		
REFERENCE	A38221		
#authors	Fuhrman, J.A.; Lane, W.S.; Smith, R.F.; Plessens, W.F.;		
#journal	Petler, F.B.		
#title	Proc. Natl. Acad. Sci. U.S.A. (1992) 89:1548-1552		
#cross-references	Transmissiion-bloking antipodes recognize microfilarial		
#accession	chitinase in brugian lymphatic filariasis.		
##status	MUID:92179220		
##molecule_type	preliminary; not compared with conceptual translation		
##residues	1-504 #label FWH		
##cross-references	NCBI:P85345		
##note	sequence extracted from NCBI backbone		
KEYWORDS	glycosidase; hydrolase		
SUMMARY	#length 504 #molecular-weight 55971 #checksum 7343		
Query Match	31.5%;	Score 1086;	DB 12; Length 504;
Best Local Similarity	42.8%;	Pred. No. 1,85e-179;	
Matches	158; Conservative	94; Mismatches 102;	Indels 15; Gaps 11;
Db	27	cyttmawgyrdgegfklpgnlpnglchlliyaafakvdeidskpfewendetskwmys	86
Qy	26	CYFTMAAYRQGEARFLPKDLDSLCHLLIYAFAGMTNHOLSTP-ENWDE-T-----LYQ	78
Db	87	avtklirenpqklyllsygyrgfsaiftgksaagcterfikaaiaflkmfdgfdl	146
Qy	79	EFNGILKKMPPKLTLLAIGMFGTQKFTDMVATNNRQTFVNSAIRLKRYSFDGLD	138



Db	147	weP-v-v-a-e-eh-kliYeamktafeeaksgkrllltaavaagkgtidsgynve	200
Oy	139	wePSSGSPAVDKKRRFTTLVQDANNAQQEELQTSIGKERLLSLAVPAGQTYVAGTEVD	198
Db	201	slgkncfdllfmsyvlbhswekncvdlhgkllhptkgevsgigifnhtefaadyaaskmpke	260
Oy	199	KIAQLLDVFNMLAVYDFHSGWEKVGHNPLPKRQEESSAAASLWVDAVQOMLOKGPAS	258
Db	261	kllslpmyagwcltdupsetaigaasrpsasaktnpaggtaasyweickylkggkety	320
Oy	259	KLILMPYGRSFTLASSDSTRVGAPAGSGTGPGEFTKEGMLAYEVCWSW-KGATKORI	317
Db	331	hgeyagaymvgdgyrygndeeitrlknkylkkygyggaifwaldfdftgkscgkypyp	380
Oy	318	ODKY-PYIFRDNOMVGDVDESEFTKYSYLYKQKGLGAMWALDLDFAQFSCNOGRYP	376
Db	381	llnaisael 389	
Oy	377	LIQTRQEL 385	
RESULT	10		
ENTRY		A56596 #type complete	
TITLE		chitinase (EC 3.2.1.14) - tobacco hornworm	
ORGANISM		formal_name Manduca sexta #common_name tobacco hornworm	
DATE		11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 03-May-1996	
ACCESSIONS		A56596	
REFERENCES		A56596	
#authors		Kramer, K.J.; Corpuz, L.; Choi, H.K.; Muthukrishnan, S.	
#journal		Insect Biochem. Mol. Biol. (1993) 23:691-701	
#title		Sequence of a cDNA and expression of the gene encoding epidermal and gut chitinases of Manduca sexta.	
#cross-references		MUID:93357793	
#accession		A56596	
##status		preliminary	
##molecule_type		mRNA	
##residues		1-554 #label KRA	
##cross-references		NCBI:136417; NCBI:136418	
##experimental_source		larvae	
##note		sequence extracted from NCBI backbone	
KEYWORDS		glycosidase; hydrolase	
SUMMARY		#length 554 #molecular_weight 62203 #checksum 4659	
Query Match		29.9%; Score 1031; DB 12; Length 554;	
Best Local Similarity		38.8%; Pred. No. 1,04e-168;	
Matches	162;	Conservative 105; Mismatches 126; Indels 25; Gaps 19;	
Db	17	vsqsdararivcyfsmwayrpgvrygtdlprvekchllsyfifgvegnsevlldpe1 76	
Oy	16	IPWGSAAFTLCYFTWMAQYRQGEARFLPKDPSICTHLIYAFAGMTN-H-Q-LST-Tew 71	
Db	77	dvdhngfrnfcslzeshbsvxfmwaavgaagsakyslmvgaqsktrmsflrsvsflky 136	
Oy	72	N-DETLYEFGNLKMNKLTLLAIGSGMGTQKFTDMVATANNROTFFVNSAIRFLRY 130	
Db	137	dfddglaweypgaaadrgsfsgdkdkiyllylqeltrf----lry-qgkwe1-taavp1an 191	
Oy	131	SFDGDLDMERFGS--GQSPAVDKERFTTLVQDLANAFQDAQTSIGKERLLLSAAVPAQ 188	
Db	192	flrmeqyhpvelcgeidlhmeyvdlrtnwagfsadvhsprlykrphdqwayeklnvdq1h 251	
Oy	189	TYVDYGEVVDKIAQLNDEVNLMAADFHGSMKEKVGHNSPLYKROEESGAASLWVDAVQ 248	
Db	252	lweeqgsrnlkvngldprygrstlsagnmnygltflnkaagsgdpearynaagfayy 311	
Oy	249	QWLQGTGASKLLILMPYGRSFTLASSSDT-RVGA--P-ATGSGTGPGEFTKEGMLAY 304	
Db	312	eictevdkdsqwkkwdegkcpayrkytgvwygeoprsvelmvnylkgqylgamtta 371	
Oy	305	EVCS--WKG-A--TKQRQDQKVPYIFPDNKNWGVDFVDESKTKVSLYKQKGLGAAWVA 359	
Db	372	ldmddfgrl-cge-knplikllhkms-sylvpphtentltpdewarppstpsdpse 426	

QY	360	LDLDDFACSCNQGRIPLIOTLRBELSLPT-LPBGTELEVPKPGQSPFEBHPSQ	416
RESULT	11	A38368 #type complete	
ENTRY		chitinase (EC 3.2.1.14) precursor - <i>Bacillus circulans</i>	
ORGANISM		#format_name <i>Bacillus circulans</i>	
DATE		28-Jun-1991 #sequence_revision 28-Jun-1991 #text_change 03-May-1996	
ACCESSIONS		A38368	
REFERENCE		A38368	
authors		Watanabe, T.; Suzuki, K.; Oyanagi, W.; Ohnishi, K.; Tanaka, H.	
#journal		J. Biol. Chem. (1990) 265:15659-15665	
#title		Gene cloning of chitinase A1 from <i>Bacillus circulans</i> WL-12 revealed its evolutionary relationship to <i>Serratia chitinase</i> and to the type III homology units of fibronectin.	
#cross-references		MUID:9036876	
#accession		A38368	
##status		preliminary	
##molecule_type		DNA	
##residues		1-699 #label WAT	
##cross-references		GB:J05599	
CLASSIFICATION		#superfamily fibronectin type III repeat homology	
KEYWORDS		glycosidase; hydrolase	
SUMMARY		#length 699 #molecular_weight 73677 #checksum 6757	
Query Match		20.3%; Score 698; DB 10; Length 699;	
Best Local Similarity		40.1%; Pred. No. 2,46e-104;	
Matches	119;	Conservative 73; Mismatches 85; Indels 20; Gaps 18;	
Db	144	qlnkktgnpmlkllisgwtw-snfsdaataaerrefansavdflkxyfdgdld	202
QY	79	EFNGLKNNPLKTLKLLAAGNNSTQKFTDVAATANNRQTFVNSAIFELKRPSTGDLDD	138
Db	203	weypvsqldgnsrpedkqnyllllakireklaaagavqk-kylltiasgasa-ya-a	260
QY	139	MEYP--GS-QG-S-PAYDKERFTLVQDLMNAFQGEAQSGKRELLLSAAVPAQGYVDA	193
Db	261	ntelakiaaivdwlnimtydfngsqwqklsahnaplndpaasaagvndantfnvaagqg	320
QY	194	GYEVDKIKQNDDPFYNLMAVDFHSGMEKVTGCHNSPL-YKROE-ESGA-A-SLVNDAVQ	249
Db	321	hldagvpaakivlsvpfyrgvwd-gcagqgn-qgyqctcgssvgytwaasfdtydean	378
QY	250	WLQKTPSKLLILMPYGRSFTLASSSDRRVGAPATGSGSPGFTREGMLAYEV-CS	308
Db	379	ylnkgyrryndcakvpylynaenkrffisyddaeasygvtaylksyglgamtwei	435
QY	309	W--K-GAIKQKIQOKVPIPRD-NQM-VEGPDVESKTKVSYLKQKGLGAMYAL	360
RESULT	12	S51591 #type complete	
ENTRY		chitinase (EC 3.2.1.14)/lysozyme (EC 3.2.1.17) p2 precursor.	
TITLE		pathogenesis-related common tobacco	
ORGANISM		#format_name <i>Nicotiana tabacum</i> #common_name common tobacco	
DATE		15-Jul-1995 #sequence_revision 08-Sep-1995 #text_change 03-May-1996	
ACCESSIONS		S51591	
REFERENCE		S51591	
authors		Heitz, T.; Segond, S.; Kaufmann, S.; Geoffroy, P.; Prasad, V.; Brunner, F.; Fritig, B.; Legrand, M.	
#journal		Mol. Gen. Genet. (1994) 245:246-254	
#title		Molecular characterization of a novel tobacco pathogenesis-related (Pr) protein: a new plant chitinase/lysozyme.	
#accession		S51591	
##molecule_type		mRNA	
##residues		1-378 #label HEI	
##cross-references		EMBL:X78325	

#accession	#experimental	source	cv.	Samsun	NN
551632					
##molecule_type	protein				
##residues	3136;87-112;252-275;282-305;337-371				
CLASSIFICATION	#superfamily	Streptomyces	chitinase	ch140	##label HEW
KEYWORDS	glycosidase;	hydrolase			
FEATURE					
1-25					
26-371					
	#domain signal	sequence	#status	predicted	#label 516\
	#product	chitinase/lysozyme	P2	#status	predicted
	MAT				#label
371-378					
SUMMARY					
	#length	378	#molecular-weight	42019	#checksum
					6689
Query Match	17.3%	Score	584;	DB	5;
Best Local Similarity	32.7%	Pred. No.	1.33e-84;		
Matches	106;	Conservative	87;	Mismatches	108;
				Indels	23;
				Gaps	18;
Db	43	midstlthfcofacadlnpsnqlispengds-fsgfctvtrkpsvxtflsisgr	101		
Oy	44	KDLDSLTCHLILVAGWT-NHOLSTEWDELTYOEFNG-LKKMKPKLTLLAI-GCW	99		
Db	102	a-dttaygimrgrpsrksfidsisrlarfgfghldidwexp-l-s-atdmtnlill	156		
Oy	100	NEGIOTKFTDMATATANNKOTFPVNSAIRFKRISFEGCLDIDMPESSQSSPAVDKRRFTTLV	159		
Db	157	newrtaimearnsgraillltaavsyprvnglnypvesvarnlmnlmaydfgpnw	216		
Oy	160	QDLNAFGQEAQTSKERLLLSAANVAGQTVVDAGYEVDKLAQNLDPVNLMAVDHGS-W	218		
Db	217	spsgnshaglfidvynh-lsgs--dg-lhawiqagvptkklilgipfygawrlvnp	270		
Oy	219	EKV-EGHNSPTRYKROESGAAASINVDAAVOQWLQKGRPAKSKLLIGMPYGRSFTLASSS	277		
Db	271	ihdlrapaagksnyga-rddgs-m-tymlridyivgsratvymatlgydcysgsnvisy	328		
Oy	278	DTRGAPATGSGTGCPFPFKEGGMALAYIVCWSKGAITKOR-IQDOK-VP-YIFRDQWVG	334		
Db	329	dtqgsvrnkvnvlykgrglllyfaw	352		
Oy	335	DDVESFTKTVSYLKOKGLGAMW	358		
RESULT	13				
ENTRY	S47133	#type	complete		
TITLE	chitinase (EC 3.2.1.14) - fungus (Trichoderma harzianum)				
ORGANISM	23-Nov-1994	#formal_name	Trichoderma harzianum		
DATE	10-May-1996	#sequence_revision	03-Aug-1995	#text_change	
ACCESSIONS	S47133				
REFERENCE	S47133				
##authors	Carsoillo, C.; Gutierrez, A.; Jimenez, B.; van Montagu, M.;				
##substation	submitted to the EMBL Data Library, June 1994				
##description	Characterization of ech-42, a Trichoderma harzianum				
##accession	S47133				
##status	preliminary				
##molecule_type	DNA				
##residues	1-424	##label	CAR		
##cross-references	EMBL:X79381				
GENETICS					
##introns	48/2; 81/2; 98/1				
CLASSIFICATION	#superfamily	Streptomyces	chitinase	ch140	
KEYWORDS	glycosidase;	hydrolase			
SUMMARY					
	#length	424	#molecular-weight	46298	#checksum
					8773
Query Match	16.9%	Score	582;	DB	5;
Best Local Similarity	36.0%	Pred. No.	2.42e-82;		
Matches	102;	Conservative	69;	Mismatches	95;
				Indels	17;
				Gaps	13;
Db	112	qlfr-lkkanrnlkwmisigwtsw-nfpaastdanrnfakfataitlmkdwfgidv	169		

QY	78	QEFNKLKKMNRKKTLLAIGWMNCTQKFTDMVATANNRQTFVNSAIRFLRKISFDCGLD	137
Db	170	dweyp-addtgatmavlllkeiirsqldayaayg-apy-yhflislaapgeph-y-sfilhm	225
QY	138	DWEYSGSGSPAVDRKERTTLVODLANAFQOEAOISGKERLLLSAAVPAGQTYVDAGYEV	197
Db	226	sdlgvdlayvnlmaydygswsvsyoghdanflapnpsnsp-yntqgalkdykxgvpa	284
QY	198	DKIAONLDFVNLMADEFSGSWEKTYGHNSPLYKKROESGGAASLNVDAAVOOWLOKCTPA	257
Db	285	sktivygmpliygraf-esteg-14gtyqsyigs-gsw--engikdy-kvlpkagatqyd	336
QY	258	SKULLIGMTYGRSFTLIASSSDTRRGAPRGATGCTGCPFTKBEGLMAYTEVCWSMGATKO--	315
Db	337	svagayysydpsskelisfdtcdmtnckvskylhnlglgsnmfw	379
QY	316	RQDQKVPIYFRDNQWGFDDVESPKTYSVLYLKKXGIGGAMWV	358
RESULT	14		
ENTRY		QJ01975	#type complete
TITLE		chitinase (EC 3.2.1.14) 1 precursor - imperfect fungus (Aphanociadium album)	
ORGANISM		#formal_name Aphanociadium album	
DATE		03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 03-May-1996	
ACCESSIONS		QJ01975	
REFERENCE		QJ01975	
#authors		Blaiseau, P. L.; Lafay, J. F.	
#journal		Gene (1992) 120:243-248	
#title		Primary structure of a chitinase-encoding gene (chl1) from the filamentous fungus Aphanociadium album: similarity to bacterial chitinases.	
#accession		QJ01975	
#molecule_type		mRNA	
##residues		1-423	#label B1A
##cross-references		GB:X64104	
COMMENT		This enzyme is essential for the degradation of insoluble chitin.	
GENETICS			
#gene		chl1	
#introns		47/2; 80/2; 97/1	
CLASSIFICATION		#superfamily Streptomyces chitinase chl40	
KEYWORDS		glycosidase; hydrolase	
FEATURE			
1-34			
35-423		#domain signal sequence #status predicted #label SIG\	
SUMMARY		#product chitinase 1 #status predicted #label MAT	
		#length 423 #molecular_weight 46072 #checksum 8802	
Query Match		16.7%; Score 576; DB 5; Length 423;	
Best Local Similarity		36.2%; Pred. No. 3,266-81;	
Matches		101; Conservative 65; Mismatches 95; Indels 18; Gaps 16;	
Db	115	lkkgrmmkmlslygwtwst-nfpaassaatlctkfagsvvgfmkdwfgididweyp	173
QY	83	LKKMNPKLTLLAIGGMWFGTOKFTDMAVATNNRQTFVNSAIRFLRKYSFEDGLDWEYP	142
Db	174	-adatagqnmvlllqavrseldsyaayg-akg-hbflislaapgdghnyklkfa-e-1g	228
QY	143	GSQGSPPAVDRKERTTLVODLANAFQOEAOISGKERLLLSAAVPAG-OTTYDAGYEVKIA	201
Db	229	kwldylnlmaydyagswnytcghdanlyanpnatp-yaiddavgaylngvpanklv	287
QY	202	QNLDPVNLMAVDHFSHWSKRYGHNSPLYKKROESGGAASLNVDAAVOOWLOKGPASKLI	261
Db	288	lpmpliyrsfq-qt--eg-1gkpyngljs-gsw--engikdykalpr-agatvcdclt	339
QY	262	LCMPYGRSFTLIASSSDTRVGAPRGATGCTGCPFTKBEGLMAYTEVCWSMGATKORIDOK	321
Db	340	gcysydpstkelisfdtcdmtnckvskylhnlglgsnmfw	378
QY	322	VPIYF-RD-NQWGFDDVESPKTYSVLYLKKXGIGGAMWV	358







```

Db 68 ttwendvlylgymlntlkhnropnlrtllsvgwntgsgqrfskiasntcstrrtflksvppl 127
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 68 tTtENNDtELLYEeFGKkMNPkRLtLLlAIGGwNGTOKtFDMAwTANNRtFNtSNAIRfL 127

Db 128 rthbfdfgdlaawlparr-----dkqhfntllikemkaefkeaq-pqkkgllisaaisag 181
| : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 128 RKfSPeGDDLMtEPfSGSGSAVvKkERTtIvQDLAAfQEOEAOTsKtRKLlLSAAvPAG 187

Db 182 kvtlssydiaksqhlidsslmtyddfhgawrgtltghspflfrrqedaspdrtfscntcav 241
| : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 188 QTYVADAGEvDKiNOndFvLMAvDFHGSExKtTGhNSpLYKkNOESGAASlNDVAav 247

Db 242 gylmrljgapasklvmyjptcfgrsfllass-etvgvabdispgyljgrfctkeagtlayyeic 300
| : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 248 QOWlOKGpAsPKLlLGmTtYGRStFLASSpDTRVGAfATGSGfPPTKGGMlAYEvc 307

Db 301 dflrgatvhrcllggqvypractqngvwyddqesxkskvyrlkdtqqlagmwvaldidfq 360
| : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 308 SW-KKATQRlQODKvPPIfFDNOMwGvDFDESPrKtVSLKOGKLGGAwMALDDfFA 366

Db 361 gsfcggdrlrfltnaikadala 381
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 367 GFSCNQG-RYPLlQITLRQELS 386

```

RESULT	2	STANDARD;	PRT;	504 AA.
ID	CHIT. BRUMA			
AC	P29030;			
DT	01-DEC-1992 (REL. 24, CREATED)			
DT	01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)			
DT	01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)			
DE	ENOCCHITINASE PRECURSOR (EC 3.2.1.14) (MFI ANTIGEN).			
OC	BIRGIA MALAY.			
CC	EUKAROTA. METAZOA. ACCELLOMATES; NEMATODA; SECERNENTEA; SPIRURIDA.			
RN	[1]			
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RX	MEDLINE; 92119220.			
RA	FURHRAN J.A., LANE W.S., SMITH R.F., PIESSENS W.F., PERLER F.B.;			
RL	PROC. NATL. ACAD. SCI. U.S.A. 89:11548-11552(1992).			
CC	-1- FUNCTION TO DEGRADE CHITIN ANTIGEN IS A MICROFILARIAE CHITINASE, WHICH MAY			
CC	FUNCTION TO DEGRADE CHITIN-CONTAINING STRUCTURES IN THE MICRO-			
CC	FLARIA OR IN ITS MOSQUITO VECTOR DURING PARASITE DEVELOPMENT AND			
CC	TRANSMISSION.			
CC	-1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE BETA-1,4 LINKAGES OF			
CC	N-ACERYL-D-GLUCOSAMINE POLYMERS OF CHITIN.			
CC	-1- PTM: O-GLYCOSYLATED.			
CC	-1- KNOWN TO BIND CALCIUM.			
CC	-1- DEVELOPMENTAL STAGE: THE APPEARANCE OF THE MFI ANTIGEN CORRESPOND			
CC	WITH THE ONSET OF THE PARASITE'S ABILITY TO INFECT THE MOSQUITO.			
CC	-1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL			
CC	HYDROLASES).			
DR	EMBL; M73689; G156064; -.			
DR	PIR; A38221; A38221.			
DR	PROSITE; PS01095; CHITINASE_18.			
KW	HYDROLASE; GLYCOSIDASE; CHITIN DEGRADATION; SIGNAL; ANTIGEN;			
KW	REPEAT; GLYCOPROTEIN; CALCIUM-BINDING.			
KM	REPEAT			
FT	SIGNAL	1	22	
FT	CHAIN	23	504	ENDOCCHITINASE.
FT	DOMAIN	23	400	CATALYTIC.
FT	DOMAIN	401	450	SER/THR-RICH (LINKER).
FT	DOMAIN	407	448	3 X 14 AA APPROXIMATE TANDEM REPEATS.
FT	ACT_SITE	148	148	PROTON DONOR (BY SIMILARITY).
SO	SEQUENCE	504 AA;	55971 MW;	4DAVE5E8 GRC32;

[illegible]

Db 87 avtkirenpbplkvllsvgyvgnfsaiflgtakbaqgterficsaalaflrkmmfdgld 146  
 QY 79 EPNGILKKNPKIKTLLAIGWNNFTQKFTDVAFAANNKQTFVNSAIFELKKRYSPDGDLD 138  
 Db 147 weyp-v-gv-a-e-eh-aklveanktafveeakcsqgrillltaavsaqgktldgsyve 200  
 QY 139 WEYPSGOSPAVDKERNFTTLVQDLANAFQQAQTSCKGERLLLSAAVYAGQTYDAGYEVD 198  
 Db 201 slgxnfdllfmsyvgldhgswekvnldngklhptkgevsgisgiltetfaadyasakgmpe 260  
 QY 199 KIAQNLDFVNLMAADVHGSWEKVTGHNSPLYKREESGAASAALVDAAYQOMQKGPAS 258  
 Db 261 kllsldipmyaagwtclndpsetaigaasrpsaaskltpaggtasyeickylkegketyv 320  
 QY 259 KLLCMPTPGYGSFTLASSSDRVPAGAPATGSGTPEPFTMEGGLMLYEVCSW-KGATQRI 317  
 Db 321 hgegvgaymvnvgddqvygyduneetlrlkmkwtkexgyggaftawldfdfdfgksesgkpy 380  
 QY 318 ODQK-V-PYIFRDNQWVGDDVESFRTKVSYKQKGLGAGAMWALDDDFAGFSCNQGRP 376  
 Db 381 llnaisael 389  
 QY 377 LIQITRQEL 385

ID	RESULT	3	CHIT MANSE	STANDARD:	PRF:	554	AA.
AC	P36362:						
DT	01-JUN-1994	(REL. 29,	CREATED)				
DT	01-JUN-1994	(REL. 29,	LAST SEQUENCE UPDATE)				
DT	01-FEB-1995	(REL. 31,	LAST ANNOTATION UPDATE)				
DE	ENOCCHITINASE	PRECURSOR	(EC 3.2.1.14).				
OS	MANDUCA SEXTA	(TOBACCO HAWKMOH)	(TOBACCO HAWKMOH).				
OC	EURYARCTA:	METAZOA:	ARTHROPODA:	INSECTA:	LEPIDOPTERA.		
RM	[1]						
RP	SEQUENCE FROM N.A.						
RX	MEDLINE:	93357793.					
RA	KRAEMER K.J.,	CORPUZ L.,	CHOI H.K.,	MUTHUKRISHNAN S.:			
RL	INSECT BIOCHEM.	MOL. BIOL.	23:691-701(1993)				
CC	-1- FUNCTION:	DIGEST CHITIN	IN THE EXOSKELETON	DURING THE MOLTING			
CC	PROCESS.						
CC	-1- CATALYTIC ACTIVITY:	HYDROLYSIS OF THE BETA-1,4	LINKAGES OF				
CC	N-ACETYL-D-GUCCOSAMINE	POLYMERS OF CHITIN.					
CC	-1- DEVELOPMENTAL STAGE:	HIGH LEVELS SEEN	IN THE EPIDERMIS ON DAY 0,				
CC	BUT RAPIDLY DISAPPEARS	AND IS UNDETECTED	ON DAYS 1-4 OF FIFTH				
CC	INSTAR. IT REAPPEARS	ON DAY 5 AND PEAKS	ON DAY 7 AFTER WHICH A				
CC	RAPID DECLINE IS SEEN.	IN THE GUT IS DETECTED	ON DAY 6 WITH LOWER				
CC	LEVELS SEEN ON DAYS 0, 7	AND 8.					
CC	-1- TISSUE SPECIFICITY:	EPIDERMIS AND GUT.					
CC	-1- SUBCELLULAR LOCATION:	SECRETED.					
CC	-1- SIMILARITY:	BELONGS TO CHITINASE	CLASS II (FAMILY 18 OF GLYCOSYL				
CC	HYDROLASSES).						
DR	EMBL:	U02270:	G406049, -				
DR	EMBL:	564757:	-; NOT_ANNOTATED_CDS.				
DR	HSSB:	P07254:	1CTN.				
DR	PROSITE:	PS01095:	CHITINASE.18.				
KM	HYDROLASE:	GLYCOSIDASE:	CHITIN DEGRADATION:	SIGNAL:	GLYCOPROTEIN.		
FT	SIGNAL	1	19	POTENTIAL.			
FT	CHAIN	20	554	ENDOCCHITINASE.			
FT	DOMAIN	396	453	SER/THR-RICH.			
FT	ACT_SITE	146	146	PROTON DONOR (BY SIMILARITY).			
FT	CARBOHYD	85	85	POTENTIAL.			
FT	CARBOHYD	303	303	POTENTIAL.			
FT	CARBOHYD	407	407	POTENTIAL.			
FT	CARBOHYD	545	545	POTENTIAL.			
SO	SEQUENCE	554	AA:	62203	MM:	FA878AD	CRC32:

```

Query Match      29.9%; Score 1031; DB 2; Length 554;
Best Local Similarity 38.8%; Pred. No. 1,18e-213;
Matches 162; Conservative 105; Mismatches 126; Indels 25; Gaps 19;

Db 17 vgsdrrarlvcyfsmwayrpyvgrgfyedlpvckchllsyflgvtgnsewllldpel 76
      | | : : | | | | | | | | | | : : | | | | | | | | : : |

```

```

OY 16 IPWMSAAKLVCFYFTNMAQYRQGEARFLPKDDLSLCTHLIYAFAGMTN-H-Q-LST-TW 71
Db 77 dvdkngfrfntslrshpsvfkfmaavggaaegskysghmaaqstmsftrsvsflkky 136
OY 72 N-BETLVOENGKAKKMPKRTLLALIGWMNFGTQKFTDMVATANNRQTFVNSAIRFLRKX 130
Db 137 dfgdldwvpaadagsgsfdsdkflylvgeflra---lrv-gkwel-taavplan 191
OY 131 SFGGLDMDWYFPGS--QGSFVNDKERFTTLVQDLANAFQDEAQTSGKERLLLSAAYVAGQ 188
Db 192 firmeghyhpelcqeidaahvmsydlirgnwagfadvhsplyktrphdqwayeklnvndglh 251
OY 189 TYVDAGYEVDKIAQNLDVNLMAVDFHGSWEKVTGHNSPLYKKEESGAASLNVDAVQ 248
Db 252 lweekgpcsnklylvgrfygrstlsegannnylgctfinkkaagggdapytnatqfauy 311
OY 249 QWLQKGPASKLILGMPTYSRFTLASSSDT-RVGA--P-ATGSGTGPPTKEGMLAYV 304
Db 312 elctevdkdsgvttkwwdegskcgyaykgtvgyvedgsrvelkmmwlkkylgamta 371
OY 305 EVCS--WKG-A--TKQRIQOKVPIYIRDNQWGFDDVESFKTKVSYLKKGLGAGAVMA 359
Db 372 ldmddfgql-cge-knpilk1k1hkms-sylvpphntentpewarpsdpsdse 426
OY 360 LDLDDFAGFCGCGRVPILQTLRQELSLPY-LPSGPELVPKPGQSPDEHPSPEQ 416

RESULT 4
ID CH11-BACCI STANDARD: PRT: 699 AA.
AC P20533;
DT 01-FEB-1991 (REL. 17, CREATED)
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT 01-FEB-1991 (REL. 31, LAST ANNOTATION UPDATE)
DE CHITINASE A1 PRECURSOR (EC 3.2.1.14).
GN CH11.
OS BACILLUS CIRCULANS.
OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WL-12;
RA MEDLINE: 90368776.
RA WATANABE T., SUZUKI K., OYANAGI W., OHNISHI K., TANAKA H.;
RL J. BIOL. CHEM. 265:15659-15665(1990).
RN [2]
RP MUTAGENESIS.
RC STRAIN-WL-12;
RX MEDLINE: 93366760.
RA WATANABE T., KOHORI K., MIYASHITA K., FUJII T., SAKAI H.,
RA UCHIDA M., TANAKA H.;
RL J. BIOL. CHEM. 268:18567-18572(1993).
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE BETA-1,4 LINKAGES OF
CC N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.
CC -1- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC HYDROLASES).
DR EMBL: M57601; G142688; -.
DR PIR: A38368; A38368.
DR HSSP: P07254; ICTN.
DR PROSITE: PS01095; CHITINASE_18.
KM HYDROLASE; GLYCOSIDASE; CHITIN DEGRADATION; SIGNAL; REPEAT.
FT SIGNAL 1 41
FT CHAIN 42 699 CHITINASE A1.
FT DOMAIN 42 460 CATALYTIC.
FT DOMAIN 465 549 FIBRONECTIN TYPE-III (R-1).
FT ACT_SITE 560 644 FIBRONECTIN TYPE-III (R-2).
FT MUTAEN 200 204 PROTON DONOR (PROBABLE).
FT MUTAEN 200 200 D->N: DECREASE IN ACTIVITY.
FT MUTAEN 200 200 D->E: NO CHANGE IN ACTIVITY.
FT MUTAEN 204 204 E->D,Q: LOSS OF ACTIVITY.
SQ SEQUENCE 699 AA: 73677 MW: 5A1777CC CRC32;

```

Query Match 20.3%; Score 698; DB 2; Length 699;  
 Best Local Similarity 40.1%; Pred. No. 3.91e-132;

```

Matches 119; Conservative 73; Mismatches 85; Indels 20; Gaps 18;
Db 144 qlnklyqtpnplktlsvsgtw-snrfdsaaataetrefansavdlfkynfdgvdld 202
OY 79 EFNGLKMKNKLTLLAIGWMNFGTQKFTDMVATANNRQTFVNSAIRFLRKYSFDGDL 138
Db 203 wepvgvgldgskrpdknyelllskirekidaagavqk-kyllttsaggsatya-a 260
OY 139 WEYP--GS-QG-S-PVADKERFTTLVQDLANAFQDEAQTSGKERLLLSAAYVAGQ 193
Db 261 ntelaklaavdvmltmydfngawqklsahnaplhydpaaasaaypdantfnvaagq 320
OY 194 GYEVDRKQNLVDVNLMAVDFHGSWEKVTGHNSP-L-YKQF-ESGAA-A-SLVDAVQ 249
Db 321 hldgypaaklvlgvfygrgd-gcaagagn-ggyqctcygssvgtweagsgfdydean 378
OY 250 WLQKGPASKLILGMPTYSRFTLASSSDT-RVGA-PATGSGTGPPTKEGMLAYEV-CS 308
Db 379 ylnkgytryndatavpylynasnkrflsyddaesvgyktajlkakjlgamfwe 435
OY 309 W--K-GATKQRIQOKVPIYIRDNQWGFDDVESFKTKVSYLKKGLGAGAVMAL 360

RESULT 5
ID CH11 APHAL STANDARD: PRT: 423 AA.
AC P32470;
DT 01-OCT-1993 (REL. 27, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE CHITINASE 1 PRECURSOR (EC 3.2.1.14).
GN CH11.
OS APHANOCADIUM ALBUM.
OC EUKARYOTA; FUNGI; DEUTEROMYCOTINA (IMPERFECT FUNGI).
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ETHM 483;
RX MEDLINE: 93013040.
RA BLAISEAU P.-L., LAFAY J.-F.;
RL GENE 120:243-248(1992).
RN [2]
RP SEQUENCE OF 35-57.
RX MEDLINE: 92136437.
RA BLAISEAU P.-L., KUNZ C., GRISON R., BERTHEAU Y., BRUGO Y.;
RL CURR. GENET. 21:61-66(1992).
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE BETA-1,4 LINKAGES OF
CC N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC HYDROLASES).
DR EMBL: X64104; G429026; -.
DR PIR: J01975; J01975.
DR HSSP: P07254; ICTN.
DR PROSITE: PS01095; CHITINASE_18.
KM HYDROLASE; GLYCOSIDASE; CHITIN DEGRADATION; SIGNAL; ZYGOGEN.
FT SIGNAL 1 22 POTENTIAL.
FT PROPEP 23 34 POTENTIAL.
FT CHAIN 35 423 CHITINASE 1.
FT ACT_SITE 171 171 PROTON DONOR (BY SIMILARITY).
SQ SEQUENCE 423 AA: 46058 MW: 88456BE6 CRC32;

```

Query Match 16.7%; Score 576; DB 2; Length 423;  
 Best Local Similarity 36.2%; Pred. No. 7.00e-103;  
 Matches 101; Conservative 65; Mismatches 95; Indels 18; Gaps 16;

DB 202 QNLDNFVNLMAFYDHFHSGMEKVTGHNPSLYKROESSGAASLNVDAVQOVLQGTSPASKRI 261  
DB 288 lmplygrsfq-gt-eg-igkpyngigs-gsw--englwdykalpk-agatvkcddtak 339  
OY 262 lmplygrsfq-gt-eg-igkpyngigs-gsw--englwdykalpk-agatvkcddtak 321  
DB 340 gcyysydsckeljsfdtpamistkswlkqkglgfmfw 378  
OY 322 vpyif-rd-nqwgfdvdesfktkvsylkqkglgfmfw 358

RESULT 6  
ID CH14 TRIHA STANDARD; PRT; 423 AA.  
AC P48827;  
DT 01-FEB-1996 (REL. 33, CREATED)  
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
DE 42 KD ENDOCHITINASE PRECURSOR (EC 3.2.1.14).  
GN CHIT42.  
OS TRICHODERMA HARZIANUM.  
OC EUKARYOTA; FUNGI; DEUTEROMYCOTINA (IMPERFECT FUNGI).  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 35-52; 93-107; 371-385 & 397-414.  
RX MEDIANE; 95269313.  
RA GARCIA I., LORA J.M., LA CRUZ J., BENITEZ T., LLOBELL A.,  
PINTOR-TORO J.A.;  
RL CURR. GENET. 27:83-89(1994).  
CC -1- FUNCTION: MORPHOGENETIC ROLE DURING APICAL GROWTH, CELL DIVISION  
AND DIFFERENTIATION (CELL WALL MORPHOGENESIS). ANTI-FUNGAL  
AGENT.  
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE BETA-1,4 LINKAGES OF  
N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.  
CC -1- SUBCELLULAR LOCATION: SECRETED.  
CC -1- INDUCTION: SPECIFICALLY INDUCED BY CHITIN AND IS CATABOLITE  
REPRESSED.  
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
HYDROLASES).  
EMBL: S78423; G999376; -.  
KW HYDROLASE; GLYCOSIDASE; CHITIN DEGRADATION; SIGNAL; ZIMOGEN;  
KW CHITIN-BINDING.  
FT SIGNAL 1 22 POTENTIAL.  
FT PROPEP 23 34  
FT CHAIN 35 423 42 KD ENDOCHITINASE.  
FT ACT\_SITE 171 171 PROTON DONOR (BY SIMILARITY).  
FT CARBOHD 218 218 POTENTIAL.  
FT SEQUENCE 423 AA; 46056 MW; 899DA50A CRC32;

Query Match 16.3%; Score 561; DB 2; Length 423;  
Best Local Similarity 35.7%; Pred. No. 2,55e-99;  
Matches 101; Conservative 71; Mismatches 94; Indels 17; Gaps 13;

DB 111 qlfk-vkkanrgklvlsigvwtst-nfpsaastanknfaktitfmkdgfgidid 168  
OY 78 QENNGKKNKPKTKTLAIGMNFQTKFTDMVATANNROTFFVNSAIRLKRKTSFGLDL 137  
DB 169 dweyp-adatqsgnmlllkeyvsqrdayaay-apg-yhllitlaapsgkdy-skilr 224  
OY 138 DMFTFPGSGSPAVDKERFTTLVODLANAFQEQOTSGKRELLLSAIVPAQGVYVDAGYEV 197  
DB 225 adlgyvldylnlmaydagsfslptghdanlfmnpnnaptr-futdsavkdylingvpa 283  
OY 198 DRIAQNLDFVNLMAFYDHFHSGMEKVTGHNPSLYKROESSGAASLNVDAVQOVLQGTSP 257  
DB 284 nktvlmplygrsf--qntag--igqlyngvgs-gsw--eagldykalpk-agatvgyd 335  
OY 258 SKLILGMPYGRSFTLASSSDTRVGPATGSGTGPFTKEGMLAYEVCSSMGATKORIDOK 317  
DB 336 svakgyysnatkelsfdtpamistkswlkqkglgfmfw 378  
OY 318 QOQKVPYIFRD--Nqwgfdvdesfktkvsylkqkglgfmfw 358

RESULT 7  
ID CH11 COCIM STANDARD; PRT; 427 AA.  
AC P54196;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE ENDOCHITINASE I PRECURSOR (EC 3.2.1.14) (COMPLEMENT-FIXATION ANTIGEN)  
(CF-ANTIGEN) (CF-AG).  
GN CTSL.  
OS COCCIDIODES IMMITIS.  
OC EUKARYOTA; FUNGI; DEUTEROMYCOTINA (IMPERFECT FUNGI).  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-C735;  
RX MEDIANE; 96144270.  
RA PISHKO E.J., KIRKLAND T.N., COLE G.T.;  
RL GENE 167:173-177(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SILVEIRA;  
RA YANG C., ZHU Y., MAGEE D.M., COX R.A.;  
RL SUBMITTED (MAR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE BETA-1,4 LINKAGES OF  
N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.  
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
HYDROLASES).  
EMBL: I41663; G1200190; -.  
DR EMBL; U51271; G1256769; -.  
DR EMBL; U33265; G1255728; -.  
KW HYDROLASE; GLYCOSIDASE; CHITIN DEGRADATION; CHITIN-BINDING; SIGNAL;  
KW GLYCOPROTEIN.  
FT SIGNAL 1 2 POTENTIAL.  
FT CHAIN 2 427 ENDOCHITINASE 1.  
FT CARBOHD 387 387 POTENTIAL.  
FT CONFLICT 15 47 RMLSLRCLFCELGARFMFTLSVPAVTVTDIQ -> VOAS  
SMSSMPNYPPVPADEGGRSVYFVNW (IN REF. 2).  
FT CONFLICT 199 199 K -> N (IN REF. 2).  
FT SEQUENCE 427 AA; 47629 MW; 6CB9AF73 CRC32;

Query Match 15.9%; Score 546; DB 2; Length 427;  
Best Local Similarity 34.1%; Pred. No. 9.12e-96;  
Matches 95; Conservative 71; Mismatches 94; Indels 19; Gaps 17;

DB 115 lknknrlkllslsgwvly-spnftkfsteggrkfkfadtstlknkldgfgididwep 173  
OY 83 LKKNPKTKTLAIGMNFQTKFTDMVATANNROTFFVNSAIRLKRKTSFGLDLWEIP 142  
DB 174 ede-kqandfvlllkacreaadaysak-hpnqk-kfllitlaapagpny-nk-lklaemd 228  
OY 143 GSGSPAVDKERFTTLVODLANAFQEQOTSGKRELLLSAIVPAQ-QTYVDAGYEVDKIA 201  
DB 229 klyldfnlmaydfsgwdkvsghmsnvpfsttkp-estpfdsdkavkylagvpankly 287  
OY 202 QNLDNFVNLMAFYDHFHSGMEKVTGHNPSLYKROESSGAASLNVDAVQOVLQGTSPASKRI 261  
DB 288 lmplygrsf--ast-dg-igqlyngvgs-gsw--engvdykdmpp-qgagvteledia 339  
OY 262 LMPYGRSFTLASSSDTRVGPATGSGTGPFTKEGMLAYEVCSSMGATKORIDOK 321  
DB 340 asysdknkryllsydytklaagkaeyltnkgmggm-w 377  
OY 322 vpyif-rd-nqwg-vgfdvdesfktkvsylkqkglgfmfw 358  
RESULT 8  
ID CH14 SERMA STANDARD; PRT; 563 AA.  
AC P07254;  
DT 01-APR-1988 (REL. 07, CREATED)  
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)  
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
DE CHITINASE A PRECURSOR (EC 3.2.1.14).  
GN CHITA.  
OS SERRATIA MARCESCENS.



Query Match	12.7%	Score 437	DB 2	Length 563
Best Local Similarity	28.8%	Pred. No. 2,886-70		
Matches 93	Conservative 84	Mismatches 123	Indels 23	Gaps 19
Db	231	pfaaagkqkqvntawddpyknfnglmaikahpdklpslsgwtlsdpffmgkvx	289	
Qy	57	AFAGTNTQQLSTTEHND-T-LYGFQNLKMMNPLKTLNLIGMNFQTKFTMVATAN	114	
Db	290	-rdtfvsgvkeflgtwklfdgvdldwefpggkgnanplsgspqdeetyvlmkelramldq	348	
Qy	115	NROTFVNSAIRFLRKYP-DCGLDDWEXPGSGS-PAY---DKRFTTLQDLANFQQ	168	
Db	349	Isaeqr-kyeltaisagkdkidkvaaynvq--nsmdhflmsydfdygpfidlnlghqt	405	
Qy	169	EAQTSKERRLLISAAPVAGQTYVD-AGEYVDKIAONIDFVNILMAYDPHGSWE-KVTGHSN	226	
Db	406	alnawpkrpdcayt-tvng-vnalagvgkqkvvnvgftamgqwtvgnvqnnlptfgr	463	
Qy	227	PLTKROESSGAALSLNDAAVQOMLKQKTPSKILLGHPYTGSRFTLASSSDTVGAPAT	286	
Db	464	atg-pvktwknjngvdyrqiaqgflmsgewqytlcyataeaayvfkpftsgdlitfdarsvq	522	
Qy	287	GSSTGPGPTKGGMLAYEYCS-WKGATKQ-RIDQ-QVPIYFR-DN-QWGFDDVESFK	341	
Db	523	akgyvldkqjgglfsweldadn	545	
Qy	342	TKVSTLKQKGLGAGAWMALDDDD	364	

RESULT	9	STANDARD:	PRT:	551 AA.
AC	CHIT NPVAC			
AC	P41684;			
DT	01-NOV-1995 (REL. 32, CREATED)			
DT	01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)			
DT	01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)			
DE	PROBABLE ENDOCHITINASE PRECURSOR (EC 3.2.1.14).			
OS	AUTOGRAPH CALIFORNICA NUCLEAR POLYEDROSIS VIRUS (ACMPV).			
OC	VIRIDAE: DS-DNA ENVELOPED VIRUSES; BACULOVIRIDAE; EUBACULOVIRINAE.			
RM	(1)			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-C6;			
RX	MEDLINE: 94303173.			
RA	AYRES M.D., HOWARD S.C., KUZIO J., LOPEZ-FERBER M., POSSEE R.D.;			
RL	VIRUSOCT 202:586-605(1994).			
CC	-I- CATALYTIC ACTIVITY: HYDROLYSIS OF THE BETA-1,4 LINKAGES OF			
CC	N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.			
CC	-I- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL			
CC	HYDROLASSES).			
DR	EMBL: L22858; G559195; -.			
DR	PROSITE; PS00014; ER_TARGET.			
DR	PROSITE; PS01095; CHITINASE_18.			
KW	HYDROLASE; GLYCOSIDASE; CHITIN DEGRADATION; SIGNAL; GLYCOPROTEIN.			
FT	SIGNAL 1 17	POTENTIAL.		
FT	CHAIN 18 551	PROBABLE ENDOCHITINASE.		
FT	ACT_SITE 305 305	PROTON DONOR (BY SIMILARITY).		
FT	CARBOHD 173 173	POTENTIAL.		
FT	CARBOHD 444 444	POTENTIAL.		
FT	SITE 548 551	PREVENT SECRETION FROM ER (POTENTIAL).		
SO	SEQUENCE 551 AA; 61368 MM; BDB124DB CRC32;			
Query Match 11.5%; Score 395; DB 2; Length 551;				
Best Local Similarity 29.2%; Pred. NO. 1.25e-60;				
Matches 95; Conservative 80; Mismatches 122; Indels 28; Gaps 27.				
Db	221 pwaavqdkqkqkvsawv-ep-ykgnfgqgmaklnphlkllpslgywtlsdp-fyfnhdv 277			
Qy	57 AFACGTHNQSLSTERNDETLTYQ-EPENGIL--KM-NPKLITLLAGGNNFGQKRTDMVAT 112			
Db	278 -ekrnrvfdsyvkelfqwkffdgvdldwefpbgkqanpslgaadgdktyllllelram 336			
Qy	113 ANNROTFVNSAIRFLRKYSF-DGLDLDMEYRSGSGS-PAV-D-K-E-R-FTTIVQDLANA 165			
Db	337 lddleagc-grv-yeltsaisgydklavnya-e-aqsklqkflfmsydfkgsantdl 392			
Qy	166 FQQ-FAQISGGERLLLSAHPAG-QTYVDAGYEVDKIAQNIDFNLMAVDFHSGMEKVT- 222			
Db	393 gyqtivya-ps-wmseeilytthvavdalikgvgpnkllvgvanygrtgvtytdny 450			
Qy	223 GHNSLRYKROESGAAASLANDAVQMLQGTFRASKLITLMPHYGNSFTLASSSDTRVG 282			
Db	451 fsgtng-psgsgwedgyvdyrjqldlnnyvfyfdaaagasyfdkskgllsfdsvs 509			
Qy	283 APATSGSGPGFTKEGGMLAYEVCMSMGATKRIQDQ-KVPIYF-RDN-QWVGFDVES 339			
Db	510 vlgkxyvdrnkllgllfaweldadn 534			
Qy	340 FRTKVSYLKQKGLGAGMAYMALDLD 364			
RESULT	10	STANDARD:	PRT:	820 AA.
ID	CHIA_ALTSO			
AC	P32823;			
DT	01-OCT-1993 (REL. 27, CREATED)			
DT	01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)			
DT	01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)			
DE	CHITINASE A PRECURSOR (EC 3.2.1.14) (CHI-A).			
GN	CHIA.			
OS	ALTERONAS SP. (STRAIN O-7).			
OC	PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;			
OC	VIBRONACEAE.			
RM	(1)			



Db 425 gsdgqk-idaa-dygeaskyidwvmtcydfifgaewak-ngptaphpplaydipqgqfn 481  
Qy 184 vppgqrryagayevdkriaonldfvnlmavdfrhgwskvghnsplykroeesgaaa-sln 242  
Db 482 tadamakfkxkaypadklllgfygrgw-vgtgap-vgtatgpat-gty--eagied 536  
Qy 243 vdaavoomlqoktfrpasklllgmpttgrsftlasssdrtrgavpatsgtpgpftrkkgmala 302  
Db 537 y-kvlnscpatqtlag-tayahcgsnwsydtpratlksmdaegqglggaiffewfsg 593  
Qy 303 yvevcsmwkatarioridokvpyfrdnomwvgefddvesfrtkvsgylkqkglgamvwaldl 362  
Db 594 dtang 598  
Qy 363 ddafg 367

RESULT 13  
ID CHIT\_STRPL STANDARD: PRT: 610 AA.  
AC P11220:  
DT 01-JUL-1989 (REL. 11, CREATED)  
DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)  
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
DE CHITINASE 63 PRECURSOR (EC 3.2.1.14).  
GN CHTA.  
OS STREPTOMYCES PLICATUS.  
OC PROMAROTIA, FIRMICUTES; ACTINOMYCETALES; STREPTOMYCETACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 92192480.  
RA ROBBINS P.W., OVERBYE K., ALBRIGHT C., BENFIELD B., PERO J.;  
RL GENE 111:69-76(1992).  
RN [2]  
RP SEQUENCE OF 1-45 FROM N.A., AND SEQUENCE OF 31-45.  
RX MEDLINE: 88087127.  
RA ROBBINS P.W., ALBRIGHT C., BENFIELD B.;  
RL J. BIOL. CHEM. 263:443-447(1988).  
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE BETA-1,4 LINKAGES OF  
CC -1- N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.  
CC -1- INDUCTION: BY CHITIN.  
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
CC HYDROLASES).  
CC -1- SIMILARITY: CONTAINS A FIBRONECTIN TYPE IIT-LIKE DOMAIN.  
CC -1- SIMILARITY: CONTAINS A BACTERIAL-TYPE CELLULOSE-BINDING DOMAIN  
CC (CBD).  
DR EMBL: M82804; G153216; -.  
DR EMBL: M18397; G153209; -.  
DR PIR: A29912; A29912.  
DR PIR: JH0573; JH0573.  
DR HSSP: P07254; ICTN.  
DR PROSITE: PS00018; EF\_HAND.  
DR PROSITE: PS00561; CBD\_BACTERIAL.  
DR PROSITE: PS01095; CHITINASE\_18.  
KM HYDROLASE; GLYCOSIDASE; CHITIN DEGRADATION; CHITIN-BINDING; SIGNAL.  
FT SIGNAL  
FT CHAIN 1 30  
FT DOMAIN 35 140 CHITINASE 63.  
FT DOMAIN 148 229 CELLULOSE-BINDING.  
FT DOMAIN 236 610 FIBRONECTIN TYPE-III.  
FT ACT\_SITE 383 383 CATALYTIC.  
FT ACT\_SITE 383 383 PROTON DONOR (BY SIMILARITY).  
FT CONFLICT 3 3 F->I (IN REF. 2).  
SQ SEQUENCE 610 AA: 63974 MW: 2F5E8E35 CRC32;

Query Match 8.3%; Score 287; DB 2: length 610;  
Best Local Similarity 26.4%; Pred. No. 1.37e-36;  
Matches 78; Conservative 63; Mismatches 134; Indels 21; Gaps 16;

Db 320 fnglmlkaeyphklllyfgwtwsgg-fpdavknpaafakschdlvedprwadfdgi 378  
Qy 77 ydfevnglkmnprlktlllaigcnfctokfctdwatannrqotrvnsairlkrlys-fddgi 135  
Db 379 dldweypancgjlsdcetsapnaissmmkamraefgqdyilta-a---vtaadsgdgk-id 433

Qy 136 dldmewtpegsog-s--paydkerfttlviodlanafqoeaotskerilllsaaVPAGQTYD 192  
Db 434 aa-dygeaskyidwvmtcydfifgaewak-ngptaphspplnaydgiipqgqftadamakfk 491  
Qy 193 agdevdkiaonldfvnlmavdfrhgwskvghnsplykroeesgaaa-slnDAVQWL 251  
Db 492 skvvpadklllgfygrgw-vgtgap-vgtatgpa-gty--eagiedy-kvlnsc 545  
Qy 252 OKTFRPASKLLILMPYGRSFTLASSSDTRVGAAPATGSGTPGPFTRKKGMLAYEVCWMKG 311  
Db 546 patgtvag-tayahcgsnwsydtpratlksmdaegqglggaiffewfsgtting 599  
Qy 312 atkqriodokvpyfrdnomwvgefddvesfrtkvsgylkqkglgamvwaldlDDFAG 367

RESULT 14  
ID KTXA\_KLULA STANDARD: PRT: 1146 AA.  
AC P09805:  
DT 01-MAR-1989 (REL. 10, CREATED)  
DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)  
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
DE KILLER TOXIN ALPHA AND BETA SUBUNITS PRECURSOR (REF2 PROTEIN)  
DE KUDYEROMYCES LACTIS (YEAST).  
OG PLASMID PGKL-1.  
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CBS 2359/152;  
RA SOR F., FUKUHARA H.;  
RL CURR. GENET. 9:147-155(1985).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 84297209.  
RA STARK M.J.R., MILEHAM A.J., ROMANOS M.A., BOYD A.;  
RL NUCLEIC ACIDS RES. 12:6011-6030(1984).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 85037931.  
RA TOKUNAGA M., KAWAMURA A., HISHIMURA F.;  
RL NUCLEIC ACIDS RES. 12:7581-7597(1984).  
RN [4]  
RP IDENTIFICATION OF PROTEIN, AND SEQUENCE OF 30-44 AND 895-916.  
RX MEDLINE: 87004569.  
RA STARK M.J.R., BOYD A.;  
RL EMBL J. 5:1995-2002(1986).  
RN [5]  
RP SIMILARITY TO CHITINASE OF ALPHA-SUBUNIT.  
RX MEDLINE: 90259069.  
RA BRADSHAW H.D. JR.;  
RL NATURE 345:299-299(1990).  
RN [6]  
RP CHITINASE ACTIVITY OF ALPHA-SUBUNIT.  
RX MEDLINE: 91301161.  
RA BUTLER A.R., O'DONNELL R.W., MARTIN V.J., GOODAY G.W., STARK M.J.R.;  
RL EUR. J. BIOCHEM. 199:483-488(1991).  
CC -1- FUNCTION: THE ALPHA SUBUNIT IS A POTENT EXOCHITINASE. ALONG WITH  
CC THE BETA SUBUNIT IT PLAYS A ROLE IN THE INITIAL INTERACTION OF THE  
CC TOXIN WITH SENSITIVE CELLS AND ALLOW THE GAMMA SUBUNIT (THE ACTIVE  
CC TOXIN) TO GAIN ENTRY INTO THE CELL.  
CC -1- PTM: REF2 IS POTENTIALLY SPLIT BY MEMBRANE-BOUND BASIC AMINO ACID-  
CC SPECIFIC PEPTIDASE TO YIELD THE ALPHA AND BETA SUBUNITS.  
CC -1- SUBUNIT: THE KILLER TOXIN IS COMPOSED OF THREE SUBUNITS: ALPHA,  
CC BETA AND GAMMA.  
CC -1- SIMILARITY: LOCAL, TO OTHER CHITIN-BINDING PROTEINS.  
CC -1- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.  
DR EMBL: X07127; G2829; -.  
DR EMBL: X00762; G2844; -.  
DR EMBL: X01095; G2849; -.  
DR PIR: S07915; S07915.  
DR HSSP: P02877; IHEV.  
DR PROSITE: PS00026; CHITIN\_BINDING.



(TM)

Distribution rights by Intelligenetics, Inc

t generated

CGACGTCGGGAA

## Chapter 6

Database 0; Query 0

3333249 seqs, 555961234 bases x 2

Listing first 45 summaries

23:VIR1 24:VIR2

53: PAT2 54: PAT3

94:VRL3 95:VRL4

115:part1 116:p

mean 11.656; variance 4.533; scale 2.572

ved by analysis of the total score distribution

### Length DB ID

**AACHILA**

15, 1105 AZ, Amsterdam, The Netherlands

15, 1105 AZ, Amsterdam, The Netherlands



RESULT 2 HUMH3G 1741 bp mRNA PRI 24-JAN-1994  
LOCUS Human glycoprotein mRNA, complete cds.  
DEFINITION M80927  
ACCESSION 9348911  
NID glycoprotein.  
KEYWORDS Homo sapiens  
SOURCE Homo sapiens cDNA to mRNA.  
ORGANISM Homo sapiens  
Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1741)  
AUTHORS Hakala,B.E., White,C., and Recklies,A.D.  
TITLE Human cartilage gp-39, a major secretory product of articular  
chondrocytes and synovial cells, is a mammalian member of a  
chitinase protein family  
JOURNAL J. Biol. Chem. 268 (34), 25803-25810 (1993)  
MEDLINE 94064658  
FEATURES  
source Location/Qualifiers  
1..1741  
/organism="Homo sapiens"  
/cell\_type="chondrocyte"  
/tissue\_type="cartilage"  
1..71  
5'UTR 72..134  
sig\_peptide /codon\_start=1  
CDS 72..1223  
/note="articular 39kDa glycoprotein"  
/codon\_start=1  
/evidence="experimental"  
/product="glycoprotein"  
/db\_xref="pid:9348912"  
/translation="MGVAKASQGFVYVLLQCCSAKYKVCYITSMYSQYREDGSCFPD  
ALDRFLCHITISFANISNDHIDTEMNDVLYGMLNLTKNPNPKULLSVGWNFG  
SOREKIASNTQSRRTFKSVPEFLRTGFDLAWLYPGRDRQHTTLTKEMKAE  
FIKEAPGPKQLLSAALSAGKVTIDSSYDLAKISQHLDFISIMYDFHAWRGTTGH  
HSPLEGOEDASPDREFSNTDYGMLRGAAPSKLVNGIPFGFSFLASSETGVA  
PISGPIGRFTKREAGTLYAYEICDELGLGATVHRLLGGOVPAATKGNMGVDDDES  
KSKVOYLKDRQLAGMVAALDDIDRQSGFCGDDLFPLTMAIKDALAT"  
mat\_peptide 135..1220  
/codon\_start=1  
/product="glycoprotein"  
3'UTR 1224..1741  
polya\_signal 1715..1720  
polya\_site 1741  
BASE COUNT 416 a 510 c 447 g 368 t  
ORIGIN  
Query Match 11.0%; Score 182; DB 76; Length 1741;  
Best Local Similarity 63.8%; Pred. No. 1,14e-151;  
Matches 700; Conservative 0; Mismatches 374; Indels 24; Gaps 10;

QY 349 ATATGCTAGCCACGGCCACAAACCGTCAGACCTTTTCTCAATCGGCATCAGGTTTCTGC 408  
Db 454 gaaccatgctttagatggctgagacctgctgctgctacccttga-----ggag-- 505  
QY 409 GCAAAATACAGCTTTGACGGCCTTGACCTTGACTGGAGATACCAGAAACCCAGGGAGCC 468  
Db 506 -----agacaaacagcattttaccacctatacaaggaatgaagccgaattatata 558  
QY 469 CTGCCGTAGACAGAGCGCTTACAAACCTGTGTAGAGACTTGGCAATGCTTCACAC 528  
Db 559 aggaagccag--c--caggaagaagcagctctctcagcagcagcagctgctcgagga 615  
QY 529 AGGAAGCCCAAGCTTCAGGAGGAAGAACGCTTCTTGTAGTTCAGCGGTTCCAGCTGGCC 588  
Db 616 aggtacacattgacagcagcattgacattgccaagatatacccaaacacgtgattatata 675  
QY 589 AGACCTATGTGATGCTGTGATAGAGAGTGAGAAATCCCGAAGACCTGATTTGTCTCA 648  
Db 676 gcatctgacctcaagattttcatgagcctgctgctgctgagcaccagccatcaagctccc 735  
QY 649 ACCTTATGGCTTACGACTTTCATGGCTCTTGGAGAGAGGTCAACGGACATAACAGCCCC 708  
Db 736 tgtccagagtcagagagatgacgtcctgacagattcagacacattgactatgctgtg 795  
QY 709 TCTACAAGAGGCAAGAGAGAGAGTGTGACAGCCAGCTCAACGTGATGCTGCTGTC 768  
Db 796 ggtacatgtttagagctgagagctctgctgacagtaagctggtgtagtggcatccacacctg 855  
QY 769 AACAGTGGCTGAGAAAGGGACCCCTGACAGAGCTGATCTTGGCATGCTTACCTAG 828  
Db 856 ggaagagcttcaactctgctcttc--tgaagcty--gtttgagcccaatcagagc 912  
QY 829 GACGCTCTTACACACTGGCTCTCTCATACAGACCAAGAGTGGGGGCCACAGCGGT 888  
Db 913 cgggaattcagagccgagcttccacaaagagcagagccttccactactatgagatctg 972  
QY 889 CTGGCACTCAAGGCCCTTTCACCAAGAGAGGAGTGTGGCTTCTATGATGCTGCT 948  
Db 973 acttcctccgagagacagatcacaagaccctcgccagagagctccctatgccaaca 1032  
QY 949 CCTGGAAG-G-GG-GCCACAAACAGAAATCCAGATCAGAAAGTGCCCTACATCTTCC 1005  
Db 1033 agggcaaccagtggtgtagatcagacagcaggaagcgtaaacgaagtcagctacc 1092  
QY 1006 GGGACAAACAGTGGGTGGCTTGTATGATGTGAGAGCTTCAAAACCAAGTCACTATC 1065  
Db 1153 gctcctctgcgccag 1170  
QY 1126 GCTTCTCTGCAACACAG 1143

RESULT 3  
LOCUS HSD49835 1418 bp mRNA PRI 25-JUL-1996  
DEFINITION Human YKL-39 precursor mRNA, complete cds.  
ACCESSION U49835  
NID 91457940  
KEYWORDS human.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1418)  
AUTHORS Hu,B., Trinh,K., Figueroa,W.F. and Price,P.A.  
TITLE Isolation and sequence of a novel human chondrocyte protein related  
to mammalian members of the chitinase protein family  
JOURNAL J. Biol. Chem. 271 (32), 19415-19420 (1996)  
MEDLINE 96325055  
REFERENCE 2 (bases 1 to 1418)  
AUTHORS Price,P.A.



TITLE Direct Submission  
JOURNAL Submitted (23-FEB-1996) Paul A. Price, Biology, 0322, University of California at San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0322, USA

FEATURES Location/Qualifiers

Source 1..1418  
/organism="Homo sapiens"  
/tissue-type="articular cartilage"

CDS 36..1193  
/note="Similar to human glycoprotein encoded by Genbank Accession Number M80927, human chondroitinase precursor encoded by Genbank Accession Number U29615, human vitellin glycoprotein encoded by Genbank Accession Number U09550, and to mouse secretory protein TM-1 precursor encoded by Genbank Accession Number S27879"

/codon\_start=1  
/product="YKL-39 precursor"  
/db\_xref="PID:91457941"  
/translation="MDKSLWAGVVLVLLDGGSAKLVCFYTNMSODRPGKFTPE NIDFLSHLYSPASIENTNKVIIRKSEVWLXOTINLTKNPKIKLILSIGYLFG SKGRPMWDSSTSLRLEFINSIILFRLNNEGLDYSWTPPOKENTHRTVYLIELA FOKDPTKSTKRLILTRAGVSAKROMIDNSIOVELAKDLDEINLSEDFHGSWEPLI TGHNSPLSKWMDRPPSYNVEYAVGWIHKMPSEKVMGIEITYGHSTLAEETT VGAPASGGAAPPTSESGFLAYEICQFLKGAITRLDOOVYAVVGNOMVGDYV KSMETKYQFLKNLNLGGMVIMSDMDFTGKSCNOGPPYQAVKRSISL"

sig\_peptide 36..98  
mat\_peptide 99..1190  
polya\_site 1418  
/note="18 A nucleotides"

BASE COUNT 370 a 357 c 337 g 354 t  
ORIGIN

Query Match 9.5%; Score 157; DB 74; Length 1418;  
Best Local Similarity 62.9%; Pred. No. 1.32e-125;  
Matches 710; Conservative 0; Mismatches 391; Indels 27; Gaps 10;

D 53 ctggagcagtgtagtgctcgtctcctccaggagagatcgtctacaaactgtttg 112  
Y 44 CTGGCAGGTTTCATGTCCTGCTGATGATCCCATGGGCTGTGCTGCAAACTGCTG 103  
D 113 ctacttaaccaatcgtgccagagcagcaggaaccagaaattacaccctgagaat 172  
Y 104 CTACTTCAACCACTGGCCAGTACAGACAGGGAGGCTCGCTTCCTCCCAAGACTT 163  
D 173 tgacccttcctatgctctcattcattcattcattcattcattcattcattcatt 232  
Y 164 GACCCCAAGCCTTTGCACCCACCTCATTCAGCTTCGCTGACACCAACCAAGCT 223  
D 233 tatcataagagcaagagtgagtgatgctctacacagacacacagctcaaaacca 292  
Y 224 GAGCACACCTGATGGAATGAGAGACACTCTACCAAGGTTCAATGGCTCAAGAAAT 283  
D 293 gaatcccaactgaaatctctctgcatcattgagaggtacactgtttgttccaaagg 352  
Y 284 GAATCCCAAGTGAAGACCCCTGTAGCCATCGAGGCTGGAATTTACAGCATCAGAATT 343  
D 353 caaccctatggtgattctctcaatcaatcaatcaatcaatcaatcaatcaatct 412  
Y 344 CACAGATATGTTAGCCACGCGCAACACCGTACAGACTTTGCACTCGCCATCAGATT 403  
D 413 tctggaggaacataactgtagtgatgtagtgatgtagtgatgtagtgatgtagtgat 469  
Y 404 TCTGGCAAAATACAGCTTTGACGGCTTACCTTGACCTGGAGTACCCAGGAAGCCAGG 463  
D 470 -agaa-----a-acactcat--ttcactgtgctgattcattgagtaggaagcctt 517  
Y 464 GAGCCCTGCGGTAGACAAAGAGCGCTTACCAACCTGGTGAAGGACTTGCCAAAGCCCTT 523  
D 518 tcagaaggaactccacaataatccacaagaaggtctctctgagctggggcgatcttgc 577  
Y 524 CCAGAGGAAGCCAGACCTCAGAGGAAGGAAGAGCCCTTTCTTGAGTGCAGCGCTTCCAGC 583

D 578 agggagcaaatgattgatacaacgcatcatcaagtgtgagaactggcaaaagatctgatt 637  
Y 564 TGGGAGACCTATGATGATGCTGATACGAGGTGACAAATCCGCCAAGAACTGGATT 643  
D 638 caacaacctctgtccttgaattccatcattcattcattcattcattcattcattcatt 697  
Y 644 TGTCAACCTTATGGCTACGACTTCATGCTCTGGAGAAAG-----TCAGGGAGCA 697  
D 658 caacagccctctgagcaagggtgagcagagacagagggccaaagctcctactcaatgtga 757  
Y 658 TAAACAGCCCTCTCAAGAGAGGCAAGAGAGAGTGTGCACAGCAGCAAGCTCAAGTGA 757  
D 758 atatgctgtggtgtagtggatfacataaaggagatccatcagagagaggtgtagtggat 817  
Y 758 TGCTCTGTGCAACAGTGTGCTGCAAGAGGGAGCCCTCCAGCAAGTGTGCTTGGCAT 817  
D 818 cccacacatatggcactcctcctacacatgagctc--tg--cagaacacacgtggggccc 874  
Y 818 GCTTACCTACGAGAGCTCTTACACCTGCTCCTCATCAGACACAGAGTGGGGCCCC 877  
D 875 tgctcttgcccttgagagctgctgagcccatcagaagcttcaggcttctctgacctta 934  
Y 878 AGCCACAGGGGCTGCACTCCAGGCCCTTACCAAGAGAGAGGATGCTGGCTACTA 937  
D 935 tgaagcttgccagttccttgaaggagcagaagatcacggcctcagagatcagagcttc 994  
Y 938 TGAAGTCTGC---TCCTGGAAGGGGGCCACCAACAGAGAAATCCAGATCCAGAAAGTCC 994  
D 995 ctacgcaatcagaggaacacagtggtgtggtatgatgtatgtgaagatgtagagacca 1054  
Y 995 CTACATCTTCCGGGACACACAGTGGTGGCTTGTATGATGTGAGAGACTTCAAAACCA 1054  
D 1055 ggttcagttcctaagaattcaaaccttgaggagagccatgtagctggtctatgacatga 1114  
Y 1055 GGTCACTATCTGAAGCAGAAAGGACATGGGCGGCCATGCTGGGACATGACTTGA 1114  
D 1115 tgactcaactggcaaatcctgcaacagggccctacacctgttcca 1162  
Y 1115 TGACTTCCGGCTTCTCTCTGCAACACAGGGCGGATACCCCTCATCCA 1162

RESULT 4  
LOCUS HSU58514 1444 bp mRNA PRI 24-JUL-1996  
DEFINITION Human chitinase precursor (HUMTCHIT) mRNA, exon 1a form, complete cds.  
ACCESSION U58514  
KEYWORDS g1439565  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1434)  
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Ten, F., Trevisakis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.

TITLE The WashU-Merck EST Project  
JOURNAL Unpublished  
AUTHORS Grossman, A., Matsuyama, T., Baker, E., Waterhouse, P., Sutherland, G.R. and Mak, T.W.  
Cloning of a novel lymphoid restricted human chitinase and localization to Ip13.3

TITLE Direct Submission  
JOURNAL Submitted (17-MAY-1996) Medical Biophysics, Ontario Cancer Institute, Princess Margaret Hospital, 610 University Avenue, Toronto, Ontario M5G 2M9, Canada



FEATURES	source	location/qualifiers
5' UTR		1..1434 /organism="Homo sapiens" /map="1p13.3" /chromosome="1" /tissue_type="Soares placenta" /dev_stage="8 to 9 weeks" /clone="EST 257753"
exon		1..36 /gene="HDMTCHIT" 1..76 /gene="HDMTCHIT" /note="exon 1a; first alternatively splice exon 1; contains a Kozak consensus initiator methionine"
sig_peptide		/number=1 37..114 /gene="HDMTCHIT" 37..1209 /gene="HDMTCHIT" /note="first alternatively spliced form, see GenBank Accession Number U58515 for the second alternatively spliced form"
CDS		/codon_start=1 /product="chitinase precursor" /db_xref="PID:91439566" /translation="MGATVMDQKSLMAGVVLLLLDGGSAKLVCTFTNMSQDRQERGE KFTPEIDPLCSHLIYSFASIEENKVIIRKRSVMLYQTINSLSKTKPKILSLGISG ELXSGKGFHMVDSSTSEFINSILFLNNHFDGLDVMIVDPDQKENTHFTVLIIH ELIAPFQKDFKMSKERILLAGVSAGROMDINSYOEKLAKDIFNLSPDFHGSN EKPLITGHSNPSLSKGMDRGPSSITYNEIYAVGIYHKMPESEKVMGIPITYGHSFTLLA SAETVGPAPSGPAGPITSSGFLAYIEYICPLGAKTIRLDQDQVPAVAKGNQWNV GYDVYKSMETKVKQPLKMLNIGAMISIDMDFTGKSONGPPLVQAVKRSLSL"
exon		77..106 /gene="HDMTCHIT" /number=2 107..309 /gene="HDMTCHIT" /number=3 115..1206 /gene="HDMTCHIT" /note="first alternatively spliced form, see GenBank Accession Number U58515 for the second alternatively spliced form"
mat_peptide		/product="chitinase" 310..365 /gene="HDMTCHIT" /number=4 366..516 /gene="HDMTCHIT" /number=5 517..643 /gene="HDMTCHIT" /number=6 644..771 /gene="HDMTCHIT" /number=7 772..954 /gene="HDMTCHIT" /number=8 955..1071 /gene="HDMTCHIT" /number=9 1072..1433 /gene="HDMTCHIT" /number=10 1210..1433 /gene="HDMTCHIT"
exon		1210..1433 /gene="HDMTCHIT" 1407..1412 /gene="HDMTCHIT" 1434 /gene="HDMTCHIT" /note="32 A nucleotides"
3' UTR		360 C 342 G 357 T
polyA_signal		
polyA_site		

[illegible]

Query Match	Best Local Similarity	Score	DB	Length
Matches 707; Conservative	9.4%;	155;	74;	1500;
	62.9%;	Pred. No. 1,56e-123;		
	0;	Mismatches 390;	Indels 27;	Gaps 10;
exon	173..375	/gene="HUMTCHIT"		
	/number=3			
exon	376..431	/gene="HUMTCHIT"		
	/number=4			
exon	432..582	/gene="HUMTCHIT"		
	/number=5			
exon	583..709	/gene="HUMTCHIT"		
	/number=6			
exon	710..847	/gene="HUMTCHIT"		
	/number=7			
exon	848..1020	/gene="HUMTCHIT"		
	/number=8			
exon	1021..1137	/gene="HUMTCHIT"		
	/number=9			
exon	1138..1277	/gene="HUMTCHIT"		
	/number=10			
3' UTR	1276..1439	/gene="HUMTCHIT"		
	1278..1439	/gene="HUMTCHIT"		
exon	1473..1478	/gene="HUMTCHIT"		
polyA_signal	1473..1478	/gene="HUMTCHIT"		
polyA_site	1500	/gene="HUMTCHIT"		
	/note="18 A nucleotides"			
BASE COUNT	385 a	380 c	351 g	384 t
ORIGIN				
139	gcagggtgtagagcgtctgcgtcttcctccaggagagatctgcctccaaactggttgctac	138		
48	gcagggttcatatgcttgcgtgatatcccatggggctctctgcgaatactgctgctac	107		
199	tttaaccaactggtcccaagaccgagcaggaacccaagaaattcacccctggaatatgac	258		
108	ttcaccaactggtccgacgagacagagagggggagcgttcctccgcccaagagacttgac	167		
259	ccctctctatgctctcatctcatctatctattcattgcgcagcatcgaaacaaaggtatc	318		
168	cccgaagcctttgcacccacccctcatctatgagccttcctggcatgacacacacagctgagc	227		
319	atcaagagcaagagtgtagtcatgctctacacagacacataacagctctcaaacccaagat	378		
228	accactgagtgagatgacagagactctctaccagagattgcatgagccctgaaagagatgaa	287		
379	cccaaacctgaaattctctctgctcatctgagaggtgtaccctggttggttccaaaggttcac	438		
288	cccgaagcctgagagaccctgttagccatcgagagctggaatttcagacactcagagagttcac	347		
439	ccctatgtaggttctctctacatcagctgtggaattcattatctcatcacaatccctggtctg	498		
348	gatatgtaggacacagggccacacacacgctacagacctttgtcactcagggcattgagttctg	407		
499	aggaacctaactcttgatgactgtagtcaagctgtagcttaccag--atc-aga-aga	554		
408	cgcgaatgacactttgacagggccttgacctgagactgagagagacacacagagagggagc	467		
555	a-----a-acactcat--ttactgtgctgattcatctgagttagcagaagccttcag	603		
468	cctgcccgtgagacaaagagagcgccttcacacacccctggtacagagacttgcccaatgcttcag	527		



QY	819	CCTACTAGAGAGGCTCTTACACACGCGCTCTCTCATAGACACACAGAGTGGGGGCCCA	878						
Db	888	gcatcagagccaggaataaccagcggttcacccaagagaagatctcttccctactat	957						
QY	879	GCCACAGGGGTCTGGCACTCAGGCCCTTACCCACAGAGAGGAGTGTGCTCTACTAT	938						
Db	958	gagatctggaactcctcccaagagcccaagtcgttagacccttggccaaggtccccc	1017						
QY	939	GAAGTCGTG---CTCCGGAAGGGGGCCACCAACAGAGATCCAGATCAGAAGGTGCC	995						
Db	1018	tatgcacccaagggcaaccagtggtggtatgatgcacccaagagagcgctcaaaaacag	1077						
QY	996	TACATCTTCGGGACACACAGTGGGTGGCTTGATGATGTGGAGACCTTCAAAACCAAG	1055						
Db	1078	gcaaaagtacctaagaagcagcgacgtgctgtctatggtgtgcacccctgagat	1137						
QY	1056	GTACAGTACTGTGAAGCAGAGAGGACTGGGGGGCCATGCTGTGGGACACTGGAATTGAT	1115						
Db	1138	gacttcgggggcaactctcg	1157						
QY	1116	GACTTTCGCGGCTTCCTCG	1135						
RESULT	7	SSU1990	1733 bp	mRNA	MAM	02-FEB-1996			
LOCUS		Sus scrofa	38kDa	heparin-binding glycoprotein	mRNA, complete cds.				
DEFINITION		Sus scrofa	38kDa	heparin-binding glycoprotein	mRNA, complete cds.				
ACCESSION		U19900							
NID		9643470							
KEYWORDS		.							
SOURCE									
ORGANISM		Sus scrofa							
REFERENCE		1 (sites)							
AUTHORS		Shackleton, L.M., Mann, D.M. and Millis, A.J.							
TITLE		Identification of a 38-kDa heparin-binding glycoprotein (gp38k) in							
JOURNAL		differentiating vascular smooth muscle cells as a member of a group							
REFERENCE		of proteins associated with tissue remodeling							
AUTHORS		J. Biol. Chem. 270 (22), 13076-13083 (1995)							
TITLE		2 (bases 1 to 1733)							
JOURNAL		Millis, A.J.T.							
REFERENCE		Direct Submission							
AUTHORS		Submitted (17-JAN-1995) Albert J.T. Millis, Biological Sciences,							
TITLE		University at Albany, 1400 Washington Avenue, Albany, NY 12222, USA							
JOURNAL		Location/Qualifiers							
FEATURES		1..1733							
SOURCE		/organism="Sus scrofa"							
		/note="in lambda Zap II"							
		/clone="PB38k"							
		/tissue_type="smooth muscle"							
		/dev_stage="adult"							
		67..1218							
		/note="gp38k"							
		/codon_start=1							
		/product="38kDa heparin-binding glycoprotein"							
		/db_xref="PID:9643471"							
		/translation="MGLRYAQTFVALVLLQSCAAVKLYCYTSMQYREGDSCEPD							
		AIDPLCTHIITSPANISNNEIDLEMNVTLYDPLNTKRNPKLTLSVGNMG							
		SOREKIASNTGSRFRFKSRPPLRTDFDLAMISPGRRDKRLTIIVKEMAG							
		EYREALPGETERLLSGASAKVAIDRQDIDQISDHFISILITDFEGAMRQTTH							
		KSPLERGGDASDRPSNADIVAVSYLRLGAPANKIKVQIPFGKFTLASSTVGA							
		PASGPIGGRFTKREKILAYEICDFLQATVRRPLGQGVPAATKGNOMWGGDDDSY							
		KNKAYILKSRQLAGAMVMTLDDDFGNFCGQNLREPLTSAIKDVLAAG							
		BASE COUNT	408 a	507 c	463 g	355 t			
		ORIGIN							
Query Match		9.38; Score 154; DB 48; Length 1733;							
Best Local Similarity		62.58; Pred. No. 1.69e-122;							
Matches	687;	Conservative	0;	Mismatches	389;	Indels	24;	Gaps	9;
Db	79	gtgctcagacaggtttctgcccctggtgctctccacagatgtctgctcctaagctg	138						
QY	39	GTGGCCTGGGAGGTTTCATGTGTCCTGTGATGATCCCATGGGGCTGTGTCGAAAACTG	98						
Db	139	gttgcctactatacagaactgtgtcagttacccgaagggtgtatggagctgtccacagat	198						
QY	99	GTCTGCTACTTCAACCACTGGGCCCATTTACAGAGAGGGAGGCTGCTCTCTCCAG	158						
Db	199	gcaatgcaccccttcctcgcacccacatcatctacagcttgcacaataagaacaat	258						
QY	159	GACTTGGACCCCGACCTTTCACCCACCTCATCTACGCTTCGCTGCATGACCAACAC	218						
Db	259	gagattgcaccccttgaggtggaatgagtgtgacgtctatgcacaactgcgaacatcaag	318						
QY	219	CAGGTGACCACTAGTGTGAATGAAAGAGAGCTCTTACCAAGAGTTCAAGGCTGTAG	278						
Db	319	aacaggaaccaccaaccggaagaccctctgtctgttgaagatgaaacttggcttccaa	378						
QY	279	AAGTGAATCCCAAGCTGGAAGACCTTTAGCCATCGAGGCTGGAATTTAGACATCAG	338						
Db	379	agatttccaaaatagcctccacaactcagaatgcgaagacttcaatcaagtcgtgtca	438						
QY	339	AAGTTCACATATATGTATGTAGCCACGGCCACACACGCTGAGACCTTGTCAACTCGGCATC	398						
Db	439	ccatcctcgggaaccctatgcttctgtatgtagtgaactgaactgaactcctccctgg--gc	496						
QY	399	AGGTTCTGGCAATACAGCTTGTAGCGGCTTATACCTTGACTGTGGATCGACGGAAGC	458						
Db	497	-----gag-----agacaagcgcgcatctccacactcagtcacaagagatgaagct	543						
QY	459	CAGGGAGAGCCCTGCGGTAGAACAGAGACCGGTTCACAACTCGGTACAGATTTGGCCAT	518						
Db	544	gaatttgaagggaagcctcg--ect--ggaacagagcgcgtctcgtcgaagcgaagtcgtg	600						
QY	519	GCTTTCAGAGAGAAAGCCACAGCTCAGGAGAGAAAGCCCTTCTTGATGAGCGAGCGGT	578						
Db	601	ctctcagaggaagtcgcgcattgaagagagcgtatgaactcgcgcagatatctcaacactg	660						
QY	579	CCAGCTGGGAGACCTATGTGGAGCTTGGAATACAGGTGGACAAATCGCCCAAACTG	638						
Db	661	gacttcacagccttcttgacctatgacttccacagctgtcctgtgcgcacaactacagagac	720						

RESULT 8  
ID SSG38KD standard; RNA; MAM; 1733 BP.  
AC 247803;  
NI 9634097  
DT 19-JAN-1995 (Rel. 42, Created)  
DR 28-AUG-1996 (Rel. 49, Last updated, Version 6)  
DE S.scrofa 38kDa heparin-binding glycoprotein.  
KM gp38k; heparin-binding glycoprotein.  
OS Sus scrofa (domestic pig)  
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;  
OC Theria; Eutheria; Artiodactyla; Suidae.  
RN 1-1733  
RA 1-1733  
RT Submitted (19-JAN-1995) to the EMBL/GenBank/DBJ databases.  
RL Millis A.J., University at Albany, Biological Sciences, 1400  
RL Washington Avenue, Albany, NY, USA, 12222  
RN 1-1733  
RX MEDLINE; 95286589.  
RA Shacketon L.M., Mann D.M., Millis A.J.;  
RT "Identification of a 38-kDa heparin-binding glycoprotein (gp38k)  
RT in differentiating vascular smooth muscle cells as a member of a  
RT group of proteins associated with tissue remodeling";  
RL J. Biol. Chem. 270:13076-13083(1995).  
FH Location/Qualifiers  
FT 1..1733  
FT /organism="Sus scrofa"  
FT /clone="PBS38k"  
FT /dev\_stage="adult"  
FT /tissue\_type="smooth muscle"  
FT /clone\_1lb="lambda ZAP II"  
FT 67..1218  
FT /product="38 kDa heparin-binding glycoprotein"  
FT /db\_xref="P0634096"  
FT /translation="MGLRVAQGFVALYLLOSCAAYKLVCYTTSMQYREGDSCFPD  
A IDPFLCTHIYSFANISNNEIDTLEMDVTLTDLNLTKNRPNLKTLLSVGWNFGS  
FT R RFSKIASNTQSRRTIKSVPLPLRTHGDDGLAMISGRDRKRLTLVEMKAEFV  
FT L EALPETERLLISGAVSAGKVAIDRCYDIAIOISQHLDFISLTLYDEHGMAROTTGHSF  
FT P FRGGDASDRFSNADYAVSYVLRGADPANKLVGIPFGRSFTLASSKTDVGAPASG  
FT GIPGRFTKEKGLIAYEICDFLQGATVRRPLGQVPIYATKGNQWGYDDQESVKNKAK  
Y LKSRQLAGAMVWTLDDLDFRGNFCQNLRFPLTSAIKDVLAA"  
FT SQ Sequence 1733 BP; 408 A; 507 C; 463 G; 355 T; 0 other;  
Query Match 9.3%; Score 154; DB 15; Length 1733;  
Best Local Similarity 62.5%; Pred. No. 1,69e-122;  
Matches 687; Conservative 0; Mismatches 389; Indels 24; Gaps 9;  
Db 79 gtggtcagacagagtttggccctggtgctcagagctgtgtcgtacacagctg 138  
QY 39 gtggcctgggaggtttatggtcctgctgcatgattcccatgggctctgctgcaaaactg 98  
Db 139 gtttgctactatacagcagctgctcagtaacgggaggtgtagtggagctgtcccaagat 198  
QY 99 gtctgctacttactacacaaactggggccagtagacagagggagagctgttccctgcaag 158  
Db 199 ggcattgacccctctctcgcacccacatcatcattagctgttggcacaataagacaat 258  
QY 159 gacttgagacccagaccttggcaccaccccttacttaccgcttggcctggcagacacacac 218  
Db 259 gagattgacaccttgagtggaatgtagtgcagctcatatgacacactgacacactcaag 318  
QY 219 cagctgagacacacacagctgagatgacagagactcttaccagagattcaatggcctgaag 278

Db 319 aacaggagaccccaacactgagagacccctctgtctgttggagatggaacttggcttccaa 378  
QY 279 aagatgaatcccaagctgagacccctgtttagcctggagagctggatatttcagcactcga 338  
Db 379 agatttcccaataagcccaacactcagagctcgcaagacttcaacagtcggtgccaa 438  
QY 339 aagtttaccagatattgattaccacagggcacaacacgttcttctgaactcggccatc 398  
Db 439 ccattctcgagacccatctgttgaatgacttgagactcagctcgttgaatcccttgg-9c 496  
QY 399 agcttctctccgaataatcacagcttggacgcttgcacttgcactggagagaccaggaac 458  
Db 497 ----gag-----agacaagcgacatccacacactctgtcaagaagtgaagct 543  
QY 459 caggagagcccttccgcttaccagagagccttccacacacccctgacagacttggcccaat 518  
Db 544 gagtttgtaagggaagccctg-cc- -ggaacagagcgctcctcgtcagcgagcggtg 600  
QY 519 gcccttcacagacaggaagcccaacacctcagaggaaggaagccttcttctgagtcagcggt 578  
Db 601 tctgcaggaagagtgccatctgacagaggtatgacatcgcacagatatcacaacactg 660  
QY 579 ccagctggcagac 638  
Db 661 gacttcac 720  
QY 639 gatattgttacccttattgagcttgcacacacacacacacacacacacacacacacacac 698  
Db 721 caagtcacccctgttccgagagcgagggagtgacaggttccgagacatcagacatgagac 780  
QY 699 aacagccccccttaccagagagcagagagagagagagagagagagagagagagagagag 758  
Db 781 tatgtctgagactatgtctgagagctgaggggcccacagacacacacacacacacacacac 840  
QY 759 gctgctgacacacag 818  
Db 841 cccac 897  
QY 819 ccttacttaccag 878  
Db 898 gacatagagcag 957  
QY 879 gccacag 938  
Db 958 gagatctgagactcctcccaagagcagacagcagcagcagcagcagcagcagcagcagcag 1017  
QY 939 gaagctctg---ctcctggaagggggcagacacacacacacacacacacacacacacacac 995  
Db 1018 tatgcac 1077  
QY 996 tacatcttccggagac 1055  
Db 1078 gcaaggtacttgaagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1137  
QY 1056 gtacgctatctaacag 1115  
Db 1138 gacttcgggagacacttctg 1157  
QY 1116 gactttgccggcttctcctg 1135  
RESULT 9  
LOCUS MBRP39 1616 bp RNA ROD 28-NOV-1995  
DEFINITION M.musculus mRNA for BRP39 protein.  
ACCESSION X93035  
NID g1085065  
KEYWORDS brp39 gene; BRP39 protein.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryotes; Eutheria; Rodentia; Sclunognathl; Myomorpha; Muridae;  
Murinae; Mus.

REFERENCE	(bases 1 to 1616)
AUTHORS	Morrison, B.W. and Leder, P.
TITLE	neu and its initiate murine mammary tumors that share genetic markers generally absent in c-myc and int-2-initiated tumors
JOURNAL	Oncogene 9 (12), 3417-3426 (1994)
MEDLINE	95060797
REFERENCE	2 (bases 1 to 1616)
AUTHORS	Morrison, B.W.
TITLE	Direct Submission
JOURNAL	Submitted (14-NOV-1995) B.W. Morrison, Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA
FEATURES	Location/Qualifiers

```

/oranism="Mus musculus"
/strain="FVB/N"
/dev_stage="adult"
/tissue_type="breast"
/cell_line="SKF murine breast tumor"
/cloneline="lambda Zap II"
66..128
-----30"
misc-feature

```

```

/note="leader sequence"
66..1211
/gene="brp39"
/note="homology to g1ycosyl hydrolases"
/codon_start=1

```

BASE COUNT 420 a 442 c 404 g 350 t  
 ORIGIN

Query Match	8.0%;	Score 133;	DB 81;	Length 1616;
Best Local Similarity	62.4%;	Pred. No. 6.95e-101;		
Matches	673;	Conservative	0;	Mismatches 378;
			Indels	27;
			Gaps	13;

Db	88	caggcttgcggtctctatgcgtccagagcgctctggtataaagctgtctgact	147
Qy	49	CAGGTTTCATGCTCTCTGATGATCCCATGGGCTTGCTGCAAACTGCTGCTACT	108
Db	148	tcacagcgtgtcccgatccagtagcgggaagcgcttgaagctcttccacagacatccaac	207
Qy	109	TCACCAACTGGGCCACTACGAGAGGGGAGGCGTCCGCTTCCTGCCAAGACTTGGACC	168
Db	208	cttctcgtgtgaccacacatcatctacacgttggccaacatcagcagcaaatggtta	267
Qy	169	CCAGCCTTTGGACCCACACCTCATCTACGCGCTTCGGCTGGCATGACCAACAC -CA -GCTGA	225
Db	268	gcaatctggagtggaatcgaagcagctgcgaactatgacaaagctgataaacctcgaaccgaa	327
Qy	226	GCACCACTGATGGAATGAGACAGACTCTTACCGAGGATTCATGTGGCTGACAGACATGA	285
Db	328	acacaaacctbaaacctcctcgtcgtctgttgagggttggaatttggsgaaaaaagattc	387
Qy	286	ATCCCAAGCTGAACACCTGTTAGCCATCGGAGGCTGGAAATTCAGACTCAGAGATTCA	345
Db	388	ccgagatctccctcaaacacttagagacgcactgtcttcgctcggctgtagcccccttcc	447
Qy	346	CAGATATGATGATGCCACGCGCCAAACCGTCGACGACTTTGTCAATCGGCATCAGGTTTC	405
Db	448	tgcgtcttaatgagcttgaatgagctgtgatactcgctcgtgtcttaacct - - - - -C - - - - -	495
Qy	406	TGGCAATATACAGACTTTGAGCGGCTTGACCTTGACTGAGGATACCAGGAAGCCAGGGGA	465
Db	496	gct-taa-g-agaagaacgataattctccacccctgatacaaggaactgaatcgcggaatca	552

QY	466	GCCTCGCCGTACGACAGGACGACCTTTACACACCTGGTTACAGGACTGTGGCAATGGCTTCC	525
Db	553	caaaagaggtccag--c-cagcagagagaagaactcctgtctcagcgcagcttgtcagcag	609
QY	526	AGCGAGAGGCGCAGACCTTCAGGGAGAGAACGCGCTTCTTGATGATCGACGGTTCACAGTG	585
Db	610	gaaaagtgagccatgtgaacacgtggtatgacatgcgccagatbaaccacacccgtgatttta	669
QY	586	GCGAGACCTTATGTGGATGCTGGATACGAGGTGGACAAAATCGGCCGACACTGGGATTTTG	645
Db	670	tcaatctcatgactacatgatttccatcagagatctgcgcgcaaatcacaaggcacacaaagcc	729
QY	646	TCAACCTTATGGGCTACGACTTCCATGGGCTTTGGAGAGAAAGTCAACAGGGACATTAACAGCC	705
Db	730	ccctctcccaagggcagaaagagacatagttttgacagatacagaatgtgaactatgcg	789
QY	706	CCCTCTTCAAGAGGCAAGAGAGATGTGTGCACACCGCCGCTCAACGTGGATGTCTGTG	765
Db	790	tgcagatcatgatacgtctctggagcccaagccacagatactcagatgggcatccccaact	849
QY	766	TGCACACGTGGCTTCAGAAAGGGAGCCCTTCCACGACAGCTGATCTTGGCATGTCTACTT	825
Db	850	ttgggaagagcttcaactctgtgcatcttc-tgaaa-atcagc-tggagagcttcaatctcag	906
QY	826	ACGAGACCTCTTACACTGGCTCTCCATCAGACACACAGAGTGGGGGCCGACGACAG	885
Db	907	gggaagaatataccagagccggtttaccaagaagacagagaccctgtgcctactacagatat	966
QY	886	GGTGTGCACTCCAGGCGCCCTTACCAAGAGAGAGAGTGTGGCTCTATGAAGTGT	945
Db	967	ggaactctccaaagagagcgtgaagtcatcagctcccaagagaaggttcccttcgcta	1026
QY	946	GCTCTGTG---AAGGGGCCACCAACACAGAGATCCAGATACGAAGTGTGCCCTACATCT	1007
Db	1027	ccaagggcaaccagttggttggtgtatgtagcacaagaagagagtttcaaaaaaaggttggat	1088
QY	1003	TCCGGGCAACACCATGTGGTGGCTTTGATGATGTGACAGAGCTTCAAAACCAAGGTACAGCT	1066
Db	1087	tcctcgaagaggaagaagcttgacagagagccatgtgttgggcactggtttgttatgattt	1144
QY	1063	ATCTGAAGCAGAAAGGACTGGCGGGGCCCATGGTGTCTGGGACTGCGCTTAGATGAGACTT	1120

RESULT	10	
ID	MMND57	standard; RNA; ROD; 1506 BP.
AC	D87757;	
NI	g1545818	
DT	15-SEP-1996 (Rel. 49, Created)	
DT	15-SEP-1996 (Rel. 49, Last updated, Version 1)	
DE	House mouse; Musculus domesticus bone marrow cells mRNA for ECF-L precursor, complete cds.	
DE	ECF-L precursor.	
OS	Mus musculus (mouse)	
OC	Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia; Theria; Eutheria; Rodentia; Myomorpha; Muridae; Murinae.	
SN	[1]	
RA	1-1506	
RP	Owashi M.;	
RT	:	
RL	Submitted (06-SEP-1996) to the EMBL/Genbank/DBJ databases.	
RL	Maoto Owashi, The University of Tokushima, Faculty of Integrated Arts and Sciences; 1-1 Minami-Ohsanjiima, Tokushima, Tokushima 770.	
RL	Japan (E-mail:Owashi@ias.tokushima-u.ac.jp, Tel:0886-56-7261, Fax:0886-56-7298)	
RL	[2]	
RP	1-1506	
RA	Owashi M., Arita H., Taoka Y., Hayai N.;	
RT	"Molecular cloning and characterization of T lymphocyte-derived eosinophil chemotactic factor (ECF-L)" ;	
RT	Unpublished.	
EH	key	Location/Qualifiers
FT	source	1..1506

```

FT      /organism="Mus musculus"
FT      /cell_type="bone marrow cells"
FT      /tissue_type="bone marrow"
FT      1..1197
FT      /gene="MECF-L"
FT      /codon_start=1
FT      /product="ECF-L precursor"
FT      /db_xref="PID:g1545819"
FT      /translation="MAKLILVGLAILNLVQGLSSYQLMCYTTSNAKDRPIEGSFKPG
N      IDPCLCHLIYAFAGMKNNEITYTHEQDLRDYEAINGLKDKNTELKTLTALIGMKRFGP
A      SFSAMVSTPQNROIPIQSVIRFLROYNFDGLMDNQIPSGRSRSPKDKHLFSLVKEM
R      KAFEESEVKDIPRLLLTSTGAGIIDIVIKSGYKRIPELSQSLDYIOWMTYDLHDPRDGY
T      GENSPLYKSPYDICKSADLNDVSIISYWKDHGAASEKLIVGFPAYGHTFILLSDPSKTG
I      GAPITSTGPGKYTEDESGLAAYEVCFTFNEGATEVMDAPQEVPAAYQGENWVGIDNV
R      SFRLKAQWLKDNNGAVVWRMDMDFSGSFCHQRHFLPLTSLKGLDNLINHSACKGPY
FT      *
SQ      Sequence 1506 BP: 422 A; 337 C; 335 G; 412 T; 0 other;

Query Match      7.38; Score 121; DB 20; Length 1506;
Best Local Similarity 57.18; Pred. No. 1.14e-88;
Matches 609; Conservative 0; Mismatches 452; Indels 6; Gaps 4;

Db      70 ctgagtgtctactataccagttggctgaagagacccaataagggagttccaacct 129
Qy      96 ctggtctgtctacttccacacactggccacgtatcacagacagggagagctctcttccgcc 155
Db      130 ggtataattgacccctgctgtgtactcacctgatactatgcttcttgatgacagat 189
Qy      156 aaggaattggaccctccacgttctgacacccatctatccctctgctgacacacac 215
Db      190 aatgagatactacacacacatgagacacttgcgtgactatgaagcatgtaagtctg 249
Qy      216 caccacactggacacacactgagtggaatgagacactctctacacagagtggaatggcctg 275
Db      250 aagaagaagacacacacacacacacacacacacacacacacacacacacacacacac 309
Qy      276 aagaagaatgaatcccaagctggaagacacctgttgacacacacacacacacacacac 335
Db      310 gccctgtcaagtcacatgcttctactcctcagaacacacacacacacacacacacac 359
Qy      336 cagaagttcacagatgatgtgagacagccacacacacacacacacacacacacacac 395
Db      370 atcaagattccttctgcaataactttgagcctcaacacacacacacacacacacacac 429
Qy      396 atcaggtttctgccaatgacacacacacacacacacacacacacacacacacacacac 455
Db      430 tctcgggaagacccctcctaagacacacacacacacacacacacacacacacacacac 489
Qy      456 agccagagggagccctccctgagacacacacacacacacacacacacacacacacacac 515
Db      490 aagaacttttgaaggaatctgtggaagacacacacacacacacacacacacacacacac 549
Qy      516 aatgaccttcacagcagacacacacacacacacacacacacacacacacacacacac 575
Db      550 gtagcagggaaatctatgaatcaatgctgggtacaaagctccctgaactgctcagctc 609
Qy      576 gttccacgctgggacacacacacacacacacacacacacacacacacacacacacac 635
Db      610 ctgcaataatactcaagtcacacacacacacacacacacacacacacacacacacacac 669
Qy      636 ctggaattttgtcaacctatgagcctacgacttccatgacctcttgagagagacacacac 695
Db      670 gaaataagtcctcctataaacttcacatgacatggaagagtgctgtactcaatgtg 729
Qy      696 caataagccccctcttacaagacacacacacacacacacacacacacacacacacacac 755
Db      730 gatcaatacttccatcaggaagacacacacacacacacacacacacacacacacacacac 789

```

```

Qy      756 gatgcctgtctgtgaacagtgctgacgaagggagacccctccacgacagatgactctggcc 815
Db      790 ttccacacatgagacacacacacacacacacacacacacacacacacacacacacacac 849
Qy      816 atgccttacctagagagacgctctcactgacagtgacctctctcatcacacacacacacac 875
Db      850 cctcaataagtaactgcccacacaggaatcacacagatgacacacacacacacacacacacac 909
Qy      876 ccagccacagggctgctgacacacacacacacacacacacacacacacacacacacacac 935
Db      910 tatgaggttgatcatttttgatgaagagacacacacacacacacacacacacacacacac 969
Qy      936 tatgaagcttg--c--tcttggaag--gg--gccacacacacacacacacacacacacacac 989
Db      970 gtacccatgacctacaggtgaatgagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1029
Qy      990 gtcccttacctatcttccgggacacacacacacacacacacacacacacacacacacacac 1049
Db      1030 ttgaagctcagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1089
Qy      1050 accaagctcagctatctgacagacacacacacacacacacacacacacacacacacacac 1109
Db      1090 atgagtgactcagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1136
Qy      1110 ttgagtacttgcgcgcttctctcgaacacacacacacacacacacacacacacacacac 1156

RESULT 11
LOCUS      D87757      1506 bp      mRNA
DEFINITION House mouse; Musculus domesticus bone marrow cells for ECF-L precursor, complete cds.
ACCESSION D87757
NID      g1545818
KEYWORDS ECF-L precursor.
SOURCE      Mus musculus bone marrow cells cDNA to mRNA.
ORGANISM      Mus musculus
Eukaryote; mitochondria; eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1506)
AUTHORS      Ohashi,M.
TITLE      Direct Submission
JOURNAL      Submitted (06-SEP-1996) to the DDBJ/EMBL/GenBank databases. Makoto Ohashi, The University of Tokushima, Faculty of Integrated Arts and Sciences; 1-1 Minami-Johsanjima, Tokushima 770, Japan (E-mail:ohashi@arts.tokushima-u.ac.jp, Tel:0886-56-7261, Fax:0886-56-7298)
2 (bases 1 to 1506)
AUTHORS      Ohashi,M., Arita,H., Taoka,Y. and Hayai,N.
TITLE      Molecular cloning and characterization of T lymphocyte-derived eosinophil chemotactic factor (ECF-L)
JOURNAL      Unpublished (1996)
FEATURES
source
1..1506
location/Qualifiers
/organism="Mus musculus"
/cell_type="bone marrow cells"
/tissue_type="bone marrow"
1..1197
/gene="MECF-L"
/product="ECF-L precursor"
/db_xref="PID:g1545819"
/translation="MAKLILVGLAILNLVQGLSSYQLMCYTTSNAKDRPIEGSFKPG
NIDPCLCHLIYAFAGMKNNEITYTHEQDLRDYEAINGLKDKNTELKTLTALIGMKRFG
PASPAMVSTPQNROIPIQSVIRFLROYNFDGLMDNQIPSGRSRSPKDKHLFSLVKEM
KAFEESEVKDIPRLLLTSTGAGIIDIVIKSGYKRIPELSQSLDYIOWMTYDLHDPRDGY
GENSPLYKSPYDICKSADLNDVSIISYWKDHGAASEKLIVGFPAYGHTFILLSDPSKTG
GAPITSTGPGKYTEDESGLAAYEVCFTFNEGATEVMDAPQEVPAAYQGENWVGIDNV
SFRLKAQWLKDNNGAVVWRMDMDFSGSFCHQRHFLPLTSLKGLDNLINHSACKGPY"
BASE COUNT 422 a 337 c 335 g 412 t
ORIGIN

```







FEATURES	SOURCE
5' UTR sig_peptide CDS	
1..16 17..78 17..1636 /codon_start=1 /product="estrogen dependent oviduct protein precursor"	
1..2034 /organism="Ovis aries" /clone="MD2A, MD34, MD42" /clone_lib="sheep oviduct cDNA library of M. Murray" /sex="female" /tissue_type="oviduct"	
JOURNAL MEDLINE REFERENCE AUTHORS TITLE JOURNAL	Endocrinology 136 (6), 2485-2496 (1995) 2 (bases 1 to 2034) Desouza,M.M. Direct Submission Submitted (28-OCT-1994) Mary M. Desouza, Tufts University, Anatomy and Cell Biology, 136 Harrison Ave, Boston, MA 02111, USA Location/Qualifiers 1..2034
ORGANISM	sheep. Ovis aries Eukaryotes: mitochondrial eukaryotes: Metazoa: Chordata: Vertebrata: Eutheria: Artiodactyla: Ruminantia: Pecora: Bovidea: Bovidae: Caprine: Ovis.
REFERENCE	1 (bases 1 to 2034) Desouza,M.M. and Murray,M.K. An estrogen-dependent secretory protein, which shares identity with chitinases, is expressed in a temporally and regionally specific manner in the sheep oviduct at the time of fertilization and embryo development
LOCUS	13
DEFINITION	OVIS aries estrogen dependent oviduct protein precursor mRNA, complete cds.
ACCESSION	U16719
NID	9885600
KEYWORDS	
DB	677
QY	690
DB	677
QY	630
DB	617
QY	570
DB	557
QY	510
DB	497
QY	450
DB	437
QY	390
DB	377
QY	330
DB	317
QY	270
DB	257
QY	213

LOCUS	14	1988	1586 bp	mRNA	MAM	30-MAY-1996
DEFINITION	Ovis aries	92kD oestrus-associated oviduct glycoprotein mRNA,				
Query Match	5.7%;	Score 95;	DB 47;	Length 2034;		
Best Local Similarity	58.9%;	Pred. No. 1.15e-62;				
Matches 365;	Conservative	0;	Mismatches 252;	Indels 3;	Gaps 2;	
mat_peptide	3'UTR	polyA_signal	1637..2034	/product=estrogen dependent oviduct protein"		
	polyA_signal	1855..1860				
BASE COUNT	451 a	548 c	537 g	498 t		
ORIGIN						
	/translation="MCKLLMVLGGLMLKKHNDGAHKLVCFYFNMAASRPGSAILP #db_xref="PID:9885601" DLDPELCTHLVAFAPASNNNOIYPKPDLDKILYPERNKIKERRGKITLISGVGNF GTSITFKMLSTFSRRERFKSVIALTHTDFDGLDFLPLGLKSGPARMTVFL ELLQAFKNEAQLTMFRRLLSAIVSDPVIKADARLRLDIEIVLSYDLHGS WEKATGHNSPLESLPGDKSSAYMSYMRQLGPEKILMLPLTYGTFHLLASONE LGAAGAPASPKTKTGKAGFLAYEVCSYFORAKRINIOYPAFKGEMGYDIA ISPEYKAFKIKREHFGAMWTLLDIDFRNCPGTFPFAHLNMLNMLVDESSPS PKHREFTAVNNSRRIGPEMPMTIRDLTGLSILPLGSAVAETHKRAVITTPRGCT ATPRLPLSSGRTPAEKGTESPEKPLVSLGRLAVSPGILAVPHVLDIGKVTTP GRKNGVPEKTYIPSGKMTVDPDRAEFLERL" 79..1629 /note="secretory glycoprotein, Mr 90-92,000, flmbria and ampulla specific" /product=estrogen dependent oviduct protein"					
Db	83	aagctggtgtttatttaccacattggtgcatcagtcggccggctctgctcatctg	142			
Qy	93	AAACGGTCTCTCTACCTTCCACACCTGGGCCAGTACACAGGGGAGGCTCGCTCTG	152			
Db	143	ccccggatctgggaccccttctctgcaaccacccgtgatttgccttgctgatgac	202			
Qy	153	CCCAAGGACTTGGAGACCCAGCCTTTGACCCACCTCATCTACGCTTGTGCGCATGC	212			
Db	203	aacaatcagattgttcctaaagatccctctgataagaaatcctctacccagagttcac	262			
Qy	213	AACGACAGCTGAGAGCACCTGATGGAAATACAGAGAC-TC-TCTACAGAGATCAAT	269			
Db	263	aagctcaaggaggaagcaggagggtctgaaaaactgtctgcctggggggtgaacttc	322			
Qy	270	GGCCTGAAGAGATGAATCCCAAGCTGAAGACCTGTATGACCATCGGAGCTGGAATTTC	329			
Db	323	ggacacctggaagttcaccaaatgtctgcacagtttccaacccgggaaagttgtcaaa	382			
Qy	330	AGCACTCAGAAAGTTCCAGATATGTGTACCCAGGCCAACAAACCTGACACCTTTGTCAAC	389			
Db	383	tcaagtatcgcctcctctaagaacacatgtctctgatgcttgagacctcttctctgac	442			
Qy	390	TGCGGCATTCAGATTCTCTGCCAATACAGCTTTGACGGCTTGGACCTTGGACTGGGATAC	449			
Db	443	cctgtgactcaaggcagcccgccgctgacgcgtgagacctgtgaccttcttcttacttgaagag	502			
Qy	450	CCAGGAGCCAGGGGAGCCCTGCGGTATACAAAGAGGCGCTTCAACAACCTGTACAGAGAC	509			
Db	503	ctctctcagcgccctcaagaatgagagccagctcaacatgagcgcccaaggctctctgtct	562			
Qy	510	TTGGCCAAATGCTTTTCCACACAGGAAGCCACGACTTCAGGGAAGGAACGCTTCTCTGAGT	569			
Db	563	gctgcgcgtctctggggagcccccaagctcatcagaagaagatatgtatgacgcgtctctgggc	622			
Qy	570	GCAGCGGTTCCAGCTGGGCAACCTATGTGATGCTGGATACGAGGTGGACAAAATGCC	629			
Db	623	agaactcttgatattcataagcgtctctgcttcaagacttacaatggagctgggaaagagtc	682			
Qy	630	CAGAACCTTGGAATTTTGTAACACTTATGGCTTACGACCTTCATGGCTCTTGGAGAGAGTC	689			
Db	683	acaggacacataagccctct 702				
Qy	690	ACGGGACATAACAGCCCTCT 709				

Accession	Partial cds.
NID	U17988
KEYWORDS	9602293
SOURCE	sheep.
ORGANISM	Ovis aries
REFERENCE	Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Artiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Caprinae; Ovis.
AUTHORS	1 (bases 1 to 1596)
TITLE	Marshall,J.T., Brownlee,A.G. and Nancarrow,C.D. Cloning and sequencing of a cDNA encoding ovine oestrous glycoprotein
JOURNAL	Reprod. Fert. Dev. 8 (2), 305-310 (1996)
REFERENCE	2 (bases 1 to 1596)
AUTHORS	Marshall,J.T.A.
TITLE	Direct Submission
JOURNAL	Submitted (02-DEC-1994) James T. A. Marshall, Division of Animal Production, Commonwealth Scientific and Industrial Research Organisation, Locked Bag 1, Delivery Centre, Blacktown, New South Wales, 2148, Australia
FEATURES	Location/Qualifiers
source	1..1596
	/organism="Ovis aries"
	/strain="merino"
	/sex="female"
	/cell_type="epithelial cell"
	/tissue_type="ampulla oviduct"
	/dev_stage="oestrus"
	<1..1596
	/codon_start=1
	/product="92kd oestrus-associated oviduct glycoprotein"
	/db_xref="PID:9602294"
	/translation="LLVLKHHDAAHKLYCFETNMAFSPGASILPRDPLICTHLPVFAFNNNOIVKPDLDKELIVPEPKTKENRGAKTLISMGVNGFSRTKMLTFSPSRRFNVSVIALRTHEDFGDLEFLKRGSPARBMFVFLLELLAFKMKEAQELMRRLLSAASGSDPRVIQKATPARLGRLDPLISLSDHSMKRYTGHNSPLFSPGDKSASIASTAMSVKQDGVPEPLKGLPYGTFFHLRASQELGAGVAPPLSPGTAKGFLAYEVCSFVQAKKRWINOYVPYAFKKEWGYDAISFYKAFEEIKREFFGAMVMTLDLDIFDGNFCGTGFPVHLNINLVNDEFSSTPSPFWSTAVINSSRRPEMPMTRLDTLTGLILPIGGAVALTEHRKSAVMTTPRGATVATPRPLSGRRRAPEKTESGPEKPLSVGHLANSPPGIANVGVPHLDIGQVMPGKAKVPEKVTTPSGKATVTPDGRAETLERL"
MRNA	<1..>1596
	/product="92kd oestrus-associated oviduct glycoprotein"
	<1..36
	/note="inferred from bovine sequence: GenBank Accession Number D16639"
	/codon_start=1
	1..19
	/note="derived from PCR primer based on bovine cDNA sequence: GenBank Accession Number D16639"
	/note="derived from PCR primer based on bovine cDNA sequence: GenBank Accession Number D16639"
BASE COUNT	352 a 440 c 430 g 374 t
ORIGIN	
Query Match	5.5%; Score 91; DB 47; Length 1596;
Best Local Similarity	58.5%; Pred. No. 9,79e-59;
Matches 363; Conservativity	0; Mismatches 254; Indels 3; Gaps 2;
Db	40 aagctggtgtgtattattacaacattggagattcaagtcggccggctctcctcgatcctg 99
	93 AACACGTGTGCTACTCTACCAACATCGGCCCACTACAGACAGGGGAGGCTGCTTCC 152
	100 cccggagctgcgaccccttctctctgcgaccccaacctgtaattgccttcctcgatgaac 159
	153 CCCAAGGACTTGAGCCCGACCTTTGCACCCACACCTCATCTAGCCCTTCGCTGCAGATGCC 212
Db	160 acaatcagatgtgttctctaagatccccctcgatgagaaaatcctctaccccaagttcaac 219
	213 AACGACCACTAGACCACTGAGTGGAAAGACAGAC--TC-TCACGAGGAGGTGCAAT 269

Db	220	aagctcaagagaggaagaacagagggcgtgaataaacactgctgtccgttctgtgggggtgagaaattc	279
Qy	270	GGCCCTGAAGAAGATGAATCCCAACCTTAAGAACCCTGTATTACCATGGAGAGCTGGAAATTC	329
Db	280	ggcaactcgtgaattccacaagaatgctgtccagcttttccacaccgggaagaagtttgcacat	339
Qy	330	AGCACTCAGAGTTCCACGATATGTGTAGCCACGCGCCACACACGCTGACACCTTTGTCAAC	389
Db	340	tcaatgatgcctcctcgtgaggaacatgctgtcgaatgtcttgagcctctctcttctgac	399
Qy	390	TGCGGCATCAGAGTTCTCTCGCAATTAACAGCTTTGACGCGCCCTTGACCTTGACTGGAGTAC	449
Db	400	ccttgaaaccagagggcagagcccgcgctgacgcctgagacccttctcttacttgaagag	459
Qy	450	CCAGAACACCCAGGGGAGCCCTGCCGTGACAAAGGACGCCCTTCCAAACCTTGTTACAGAC	509
Db	460	ctcctcgagagccttcaagaatgagcccaagctcacatgctggcccaagagctgctgctgtc	519
Qy	510	TTGGCCATGCTCTTCCAGACGAGAACCCACGACCTCAAGGAAAGAAAGCCCTTCTTGAGT	569
Db	520	gtctgcgctctcttggggagacccccacgctatccagaagaacatagtcaagcctcttgag	579
Qy	570	GCAGGCGTTCCAGCTGGGCGACACTATGTGATCTGATACAGAGGTGACAAATACGCC	629
Db	580	agaccctcgattatccatagctgtctgtcttaagacttaacaggaagctggagaagatgc	639
Qy	630	CAGAACCTGGATTTTGTGTCAACCTTATGCTTACGACTTCATGCTCATGCTCTTGAGGAAAGTTC	689
Db	640	acaagagacacaatagccctctc	659
Qy	690	ACGGGACATTAACAGCCCCCT	709
RESULT	15		
LOCUS	2228 bp	mrna	PRI
DEFINITION	Papio hamadryas anubis estradiol-dependent oviduct-specific		11-OCT-1996
ACCESSION	g1256903		
KEYWORDS	estradiol-dependent oviduct-specific glycoprotein.		
SOURCE	Papio hamadryas anubis cDNA to mRNA.		
ORGANISM	Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Cercopitheciidae; Cercopitheciinae; Papio.		
REFERENCE	1 (bases 1 to 2228)		
AUTHORS	Donnelly,K.M., Fazleabas,A.T., Verhage,H.G., Mavrogianis,P.A. and Jaffe,R.C.		
TITLE	Cloning of a recombinant complementary DNA to a baboon (Papio anubis) estradiol-dependent oviduct-specific glycoprotein		
JOURNAL	Mol. Endocrinol. 5 (3), 356-364 (1991)		
MEDLINE	91367180		
REFERENCE	2 (bases 1 to 2228)		
AUTHORS	Jaffe,R.C., Atlas,E.B., O'Day-Bowman,M.B., Donnelly,K.M., Mavrogianis,P.A. and Verhage,H.G.		
TITLE	Regional Distribution and Hormonal Control of Estrogen-Dependent Oviduct Specific Glycoprotein mRNA in the Baboon (Papio anubis)		
JOURNAL	Biol. Reprod. (1996) In press		
REFERENCE	3 (bases 1 to 2228)		
AUTHORS	Jaffe,R.C.		
TITLE	Direct Submission		
JOURNAL	Submitted (05-MAR-1991) Randal C. Jaffe, Department of Physiology and Biophysics, University of Illinois College of Medicine, Chicago, IL 60612-7342, USA		
FEATURES	Location/Qualifiers		
source	1..2228		
	/organism="Papio hamadryas anubis"		
	13..1884		
CDS	/codon_start=1		
	/product="estradiol-dependent oviduct-specific glycoprotein"		
	/db_xref="PID:g1256951"		

```

/translation-"MKKLLWLVGLVLRHHDGAHKLVCYFTNMAHSRPGPASTLPH
LDPELICHILIFAPSMNNNOIVAKDLODEKILYPEPNKIKERNRELKTLISIGMNF
GTSRRTTLMSTFANREKRTIASYISLLRTHDEGDLDFLYPELRGSPMHDRTFLFLI
EELLFAFRKREALLIMRPRLLLSAAVGVPHIVQISYDVAFGLRDLFINVLSYDLGS
WEKTFGHSNPLFSLPEDEPKSSAYANMYRKLGAPESEKIMGIPYGRTEFLKASKNG
LOATNIGPASPCKYTKQAGFLAYFECISFVWGAKKHWDIOYVPYANKKEWGYDDA
ISFSYKAMFIRREHFWMTLDDVDVDFGFCGCPPLIYVMDIILVRAEFSSTL
POPLSSAVNSSNDPERLAATKATTDIKILIPGEGAGYIEIHKCNMTTPRYTI
VTPREIVSLGKHTVALGEKTEITGATMTISVGHOSMTPEKALTPVGHOSLPEKKT
LTPVGHOSVTTGOKTLISVGHSHVTPGKTLIPVGHPSVTPVSHQSVSPGGMTPVH
FOETILRONTPAIPRKAAVHEKVTVPSEKRSISVTPGQTVPLRGEYLTSETGTHPDG"
BASE COUNT      559 a      572 c      531 g      566 t
ORIGIN

```

```

Query Match      5.5%; Score 91; DB 67; Length 2228;
Best Local Similarity 58.5%; Pred. No. 9.79e-59;
Matches 363; Conservative 0; Mismatches 254; Indels 3; Gaps 1;

Db 79 aaactggtgtgtattatcccaactgggcacacagtcgagcccgccctcgatctg 138
   |||||  ||  |||||  |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||
QY 93 AACTGGTGTGCTACTTCAACACTGGGCCAGTACACAGGGGAGGCTCGCTTCG 152
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

Db 139 ccccatgacctgacccctctctcgaccacccatgatatctgacctgccaatgaac 198
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 153 CCCAAGGACTTGGACCCCGACGCTTTCACACCACCTCATCTAGCGCTTCGCTGGCATGACC 212
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

Db 199 aacaatcagattgttgtaaggatctccagagcagagaaattctaccacagatcacaac 258
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 213 AACCAACAGCTAGCACACACTGAGTGAATGACGAGAC--TCTTACACAGGAGTCAAT 269
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

Db 259 aaactaaagagagagagagagagctgaaacacactgcttccaatcgcggtggaactt 318
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 270 GGCCCTGAAGAAGATGATCCCAAGCTGTAAGACCTGTAGCCATCGAGGCTGGAATTTC 329
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

Db 319 ggcacctcagggtccacactatgtgtccacattggccaacggtgaaagttaattgtct 378
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 330 AGCACTCAGAAAGTTCCACATATGTAGTACGAGGCCAACACCGTTCAGACTTTGTCAAC 389
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

Db 379 tcaagtataatcccttcgagacacatagaacttgatgcttgaatctttttcttataat 438
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 390 TCGGCATCAGGTTCTGTGCGCAATACAGCTTTGACGGCTTGACCTTGACCTGGAGATAC 449
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

Db 439 cctgactaaggagcagcccatgcatgacgggtggaacttctctcttaattgaagag 498
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 450 CCAGGAAGCCAGGGAGCCCTCGCTGAGACAAGAGAGCGCTTCACAACTGTGTACAGGAC 509
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

Db 499 ctccctgttgccttcggaagagagcagctgcaaccatgcgcccaagggtgctgctct 558
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 510 TTGGCCAAATGCTTCACAGAGGAGCCAGACTTCAGGSAAGAACGCTTCTTCTGAGT 569
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

Db 559 gctgctgttctctgagggtccacacacatcgiccaaacatcctatgatlgtgccttctaaga 618
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 570 GCAGCGGTTCCAGCTGGGAGACACCTATGTGATGTGATACGAGGTGACAAATAATCGCC 629
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

Db 619 agactccttgatttcataatgctgtcttatgacttcatagaaattgggaaagtcc 678
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 630 CAGAACTGTGATTTGTCAACCTTATAGGCTACGACTTCATAGGCTTGGGAGAGGTC 689
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

Db 679 acagacataatagccccct 698
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 690 ACGGACATTAACAGCCCTT 709
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

Search completed: Sun Jun 29 20:45:12 1997  
 Job time : 1558 secs.





QY 109 TCACCAACTGGGCCAGTACAGACAGAGGGAGGCTGCTTCCTCCGCCAAGACTTGAGACC 168  
 Db 214 gcttcctgtgtaccacatcatcacttaagctttgccaataataagcaacgcttcacatgaca 273  
 QY 169 CCAAGCTTTGCAACCCACCTCATCTACGCTTCGCTGGCATGACCAACCAACGACTGAGA 228  
 Db 274 cctggaggtgagatgtatgtcgtcttaagcgtcgtcctcaacaacacacacacgagacc 333  
 QY 229 CCACGTGAGTGAATGACGAGACTCTTACAGAGATTCAATGGCCCGAAGAAATGATAATC 288  
 Db 334 ccaacctgaagactctctgtctgtctgagagatgaaacttgggtctcaagaatttcca 393  
 QY 289 CCAAGCTGAGAGCCCTGTGATGGCATGGAGGCTGGAATTTCAACATCAGAGATTACAG 348  
 Db 394 agatagccccaacacacacagagatgctgagacttctcaagtaagcagtaacgcaattctgc 453  
 QY 349 ATATGTTACCCACGGCCAAACCGTCAACCTTTGTCAACTGGCCATCAGATTCTCCG 408  
 Db 454 gacaccatgtcttgaatggcgctgaacctgtcctgtcctcaaccttga---c---gagag-- 505  
 QY 409 GCAAAATACAGCTTTGACGGCCTTGACCTTGACTGGAGATACCGAGAACCAAGGGAGGCC 468  
 Db 506 -----agacaaacacacatttaccacccctaatcaagaaatgaagcgcgaattataa 558  
 QY 469 CTGCCGTAGACAGAGCGCTTCACAACTGTGTACAGACTTGCCCAATGCTTCACAGC 528  
 Db 559 aggaagacccag--c--cagggaaanaagcagctcctgtctgagcgagacgtctgctggga 615  
 QY 529 AGGAAGCCCAAGACTTCAGGGAGAGAGAGCGCTTTCTGTGATGAGGAGGTTCCAGCTGGGC 588  
 Db 616 aggtcaccaatgacagcaactatgacatgtgcacagatatccacacacacacgttcatla 675  
 QY 589 AGACCTATGTGATGTGTGATAGAGAGTGTGACAAATTCGCCAGACCTGGATTGTGTA 648  
 Db 676 gcatctgacctcaagatttcatgtgcctgtgcgtgtgagacacagagccatcagatccccc 735  
 QY 649 ACCCTATGTGCTACGACTTCCTCATGTGGTGTGGAGAGGTCAAGGAGCAATACAGCCCC 708  
 Db 736 tcttcagagtgtaagagatgtcagatcctgtacagatctcagacacactgtactagctgtg 795  
 QY 709 TCTACAGAGGCAAGAGAGAGTGTGACAGACCCAGCTCAACGTGATGCTGTCTGTGC 768  
 Db 796 ggtacatgtgtgagctgtggtcctcctgcagtaagctgtgtatgagacatcccccacttcg 855  
 QY 769 AACAGTGGCTGAGAGAGGGGACCCCTGACAGAACTGATCCCTGGCATGCTCACTAGG 828  
 Db 856 ggaagagcttcaactctgtgcttctc--tgagactg--gtgttccagcgccaatctcagagc 912  
 QY 829 GACGCTCTTCAACACTGAGGCTCTCTCATCAGACACAGAGTGGGGGCCCCAGCCAGGGT 888  
 Db 913 cgggaattccagagcggttcacaaagagagcaagagaccccttgctactatgagatctgtg 972  
 QY 889 CTGGCACTCAAGGCCCTTTCACCAAGAGAGAGGAGTGTGCTGCTTATGAAGTGTGCT 948  
 Db 973 acttccctcgcgagccacagtcacatagaaacctgcgacgagcctccatgccaaca 1032  
 QY 949 CCGTGAAG-G-66-GCCACCAAAACAGAGATCCAGATCAGAAAGTGCCTCATATTTCC 1005  
 Db 1033 agggacacacagtggtgagatagacagacagcaagaaagcgltcaaaagcaagtgtagtacc 1092  
 QY 1006 GCGACACACAGTGGGTGCTTGTATGATGTGGAGAGCTTCAAAACCAAGGTACAGCTATC 1065  
 Db 1093 taaagagatgagcagctgtgcagcgcgcatgtgtatggccttgagacccggatgattccag 1152  
 QY 1066 TGAAGCAGAGAGGACTGGCGGGGCCATGTGTGGCTACTGACTTAGATGACTTTGCCG 1125  
 Db 1153 gctccttctgcgagcag 1170  
 QY 1126 GCTTCTCTGCAACGAG 1143

RESULT 2  
 ID N81756 standard; DNA; 966 BP.

AC N81756;  
 DE 19-OCT-1990 (first entry)  
 DE Gene encoding polypeptide involved in protective mechanisms  
 KW Immune response; cell growth; ss.  
 FH Key  
 FT CDS  
 FT /tag- a  
 FT PN J63032898-A.  
 PD 30-JAN-1988.  
 PF 16-JUL-1986; 167518.  
 PR 16-JUL-1986; JP-167518.  
 PA (DAIN) Dainippon Pharm KK.  
 DR WPI: 88-068419/10.  
 DR P-PSDB: P81342.  
 PT New polypeptide and DNA encoding it - related to protective mechanisms  
 PT such as immune response etc.  
 PS Disclosure: P; Japanese.  
 CC Polypeptide encoded is involved in protective mechanisms such as immune  
 CC response, cell growth and activation of protective functions.  
 SQ Sequence 966 BP; 238 A; 282 C; 247 G; 199 T;

Query Match 7.9%; Score 131; DB 1; Length 966;  
 Best Local Similarity 66.1%; Pred. No. 2,21e-69;  
 Matches 269; Conservative 0; Mismatches 138; Indels 0; Gaps 0;

Db 23 caggtctgtgtctgtctgtgtcgtcagtgctgtgtctgtcacaactgtgtctact 82  
 QY 49 CAGGTTTATGTTGCTCTGCTGATGATCCATGGGGCTGTGCTCCAAACCTGCTCTACT 108  
 Db 83 acaccagctgtgccagtaacgagagcgatgagagctgtcctccagatgacctgacc 142  
 QY 109 TCACCAACTGGGCCAGTACAGACAGAGGGAGGCTGCTTCCTGCCAAGACTTGAGAC 168  
 Db 143 gcttcctgtgtaccacatcatcactaagcttgcacataataagcaagatcactgcaga 202  
 QY 169 CCAAGCTTTGCAACCCACCTCATCTACGCTTCGCTGGCATGACCAACCAACGCTGACA 228  
 Db 203 cctggaggtgagatgtatgtgagcttctcagatgtctcaacacactcaagacagagacc 262  
 QY 229 CCACGTGAGTGAATGACGAGCTCTCTTCACAGAGATTCAATGGCTGAGAGATGAATC 288  
 Db 263 ccaactgaaagactctctgtctgcgagagatgaaacttgggtcctcaaaatttcca 322  
 QY 289 CCAAGCTGAAGACCTGTGATGATCGAGAGGAGGCTGGAATTTGACACTCAGATCAGACAG 348  
 Db 323 agatagcttcaacacacagagtgctgagacttctcaatgaagtagtaccgcaattctgc 382  
 QY 349 ATATGTTACCCAGGCGCAACCCCTCAGACCTTGTGCAACTCGGCCATCAGATTCTGC 408  
 Db 383 gacccatgtgtatgtatgtgtgagaccttgctgctcactaaccttga 429  
 QY 409 GCAAAATACAGCTTTGAGGGCCTTGACCTTGACTGGAGATCCACGAG 455

RESULT 3  
 ID 010572 standard; DNA; 1047 BP.  
 AC 010572;  
 DE 09-APR-1991 (first entry)  
 DE Human Natriuretic Peptide Receptor B.  
 KW NPB; ANP; BNP; CNP; kidney failure; heart failure; protein kinase;  
 KW hyperaldosteronism; glaucoma; guanyl cyclase.  
 OS Homo sapiens.  
 FH Key  
 FH Peptide  
 FT /label- signal sequence  
 FT Protein  
 FT /label- mature NPB  
 FT /label- mature NPB  
 FT Domain  
 FT /label- extracellular domain  
 FT /note- "binds natriuretic peptides A,B and C"  
 FT Domain  
 FT /label- transmembrane domain  
 FT 479..1047

FT	/label=cytoplasmic domain
FT	/note="GC and protein kinase activity"
FT	Modified -site 24..26
FT	/label=N-glycosite
FT	Modified -site 35..37
FT	/label=N-glycosite
FT	Modified -site 161..163
FT	/label=N-glycosite
FT	Modified -site 195..197
FT	/label=N-glycosite
FT	Modified -site 244..246
FT	/label=N-glycosite
FT	Modified -site 277..279
FT	/label=N-glycosite
FT	Modified -site 349..351
FT	/label=N-glycosite
FT	Modified -site 600..602
FT	/label=N-glycosite
RN	W09100292-A.
PD	10-JAN-1991.
PF	22-JUN-1990; U03586.
PR	23-JUN-1989; US-370673.
PA	(GETH ) GENENTECH INC.
PI	Chang M, Goeddel D, Lowe D;
DR	WP1: 91-036711/05.
DR	N-PDSB: Q10324.
PT	Neuriretic protein receptor B - for diagnosis and treatment of
PS	Kidney failure, heart failure, hyperaldosteronism, glaucoma etc.
PT	Claim 3, Fig 1, 49pp, English.
CC	The sequence was derived from the DNA encoding natriuretic peptide
CC	receptor B, NPRB, having guanyl cyclase (GC) activity and protein
CC	kinase activity. The DNA can be inserted into expression vectors
CC	for the prodn. of the protein, opt. after being mutated to produce
CC	NPRB analogues. The protein has a mol wt. of 115 kD (calculated Mr=
CC	114,992). The protein (or variants) can be used in treatment of
CC	natriuretic peptide disorders, and also to isolate peptides using
CC	affinity chromatography. Antibodies with affinity for NPRB can
CC	also be prepd.
SQ	Sequence 1047 BP; 87 A; 15 C; 83 G; 51 T;
 Query Match 5 6%; Score 92; DB 2; Length 1047; Best Local Similarity 10.6%; Pred. No. 1,24e-42; Matches 106; Conservative 279; Mismatches 603; Indels 15; Gaps 14;	
Db	5 snnnvnaaaagvvrnngarnltanvnnhnhsyawaawrygnavanaavangranvnd 64
OY	120 GCCCAGTCAACACAGGGGGAGCGCTCGCTCTGCCAAGACTTGACCCTTCGC 179
Db	65 trvsannngacsnyannasvdnkynhdndnngngcvynaas-varnashtwrnmntag 123
OY	180 ACCCACCTCATCTACGCGCTTCGTGGCATGACCAACCCACGCTGAGCAGATCGTAGTGG 239
Db	124 awagsnskndhyrtnvtgnsankngnvvtlnghnmwtearaannyndatrdinhnt 183
OY	240 AATGACGAG-ACCT-C-TACCAGAGAGTTCTATGCGCTGAAGAAGATGATCCCCAACCTG 296
Db	184 ngyvnanngnsnvnhvyarngnnaathnrrangrnrvncgnmmhmnnman-r 242
OY	297 AAGACCCCTTAAAGCCATGGGAGGCTGGAATTTCAGCACTCGAAGACTCACAGATATGGTA 356
Db	243 nmtngdyvnnyndvgnsnragnltrabgrvndrtlrnnaanrnaantvrvnltyrnn 302
OY	357 GCCACGGCCACAACCCGTCAGACCTTTGTCAATCGGCGCATCAGAGTTCTTGCCCAATAC 416
Db	303 nmnynmnmrmrarrndhyvrngnsmmnnaagcnydgmmnyanvnmntmngtlnrdgn 362
OY	417 AGCTTTCAGCGCCCTTGACCTTGACTGGGAGTACCCAGAAACCCAGGGAGGCC-TGCCGT 475
Db	363 rrvnkmgrriryhyvrgnvvmkknrdndrvnwmgdnsgdmnaahysganknmwttg 422
OY	476 AGACAAAGAGAGC-GCTTCACAACCTGTGTACAGGACTTGGCCAATGCTTCCACAGAGAG 534
Db	423 rrrnn-wv-kgaansdnmcandndnscktnstnaevangtgnltnmngvssnnmrk 480

OY		CCACAGACTCAGGGAAGAAGCGCTTCTTCGTGAGTGACGGGTTCCAGCTGAGCACACT	594
Db	481	mnmknksamwrrvrmnnnnnqnryhkqgsrntdnsgssygsmtbkxymnant	540
OY	595	ATGTGTGATGC-TGGATTACGAGGTGACAAATAGCCCAAGAACCTGGATTGTGCACCTT	653
Db	541	ghnkynvanvkhnvkkkrnnltrvnannkkmrdvnnhhnltnngaendnnnenvtlycn	600
OY	654	-ATGGCTTACAGACTTCCATGTGCTCTGTGGAGAAAGTCACGGGACATTAACGCCCTCTA	712
Db	601	rysmndnnndsnnmdwmrysmndrvkmmnhnsnshsnksncvdsrynvhk	660
OY	713	CAAGAGGCAAAAGAGATGGTGTGCAGAGCCAGCCTCAACGTGATGCTCTGTGCACA	772
Db	661	tldygnastrctandndnnaanyakkntcanmsngmntttgmaadyvsngnnmnar	720
OY	773	GTEGCTGCAGAAAGGGAGACCCTGCCACAGACTATCTTGGCATGCTTACTACGAGCG	832
Db	721	sgnyngndshsknvkvkrngnrnyrnndrtinnnnhvnmrcwandhanrndng	780
OY	833	CTCCTTCACTACTGGCTCCTCATGACACACAGATGGGGGCCACACAGAGTCTGG	892
Db	761	nkgmrtrnnkngltsndnnmrnmnyannnkhnvartlnaynnkrikanaannyannh	840
OY	893	CACCTCCAGCCCCCTTTCACCAAGGAAGGAGTGTGCTGCTACTATGAATGTGCTCTG	952
Db	841	sवाननकगतेवमनानदसवनतमयनतमानदवासरुनरुहभनरु	958
OY	953	GAAGGGGGCCACCAAACGAGAAATCCAGGATCAGAGTGCCTCTACATCTCCGGGCA	1010
Db	899	ndvkvvtlnddaymvsvgngrngnhanmanadvasnurhnrhndmrnrn	958
OY	1013	CCAGTGGGTGGGCTTTGATGATGAGAGAGCTTAAAAACCAAGTCAAGTCAATGGA	107
Db	959	gvhtgvcagvygnkvmrzcngdltvntlastrmsngnanxnkhv	1001
OY	1073	GAAGGGAGCTGGGGCG-GGCCATGCTC-TGGGCACCTGACACTTAG	1113
<hr/>			
RESULT	4		
ID	Q090443	standard; cDNA to mRNA; 2504 BP.	
AC	Q090443;		
DR	10-JAN-1996	(first entry)	
DE	Murine oviduct specific glycoprotein cDNA.		
KW	Murine oviduct specific glycoprotein; recombinant production; MOGP; ds.		
OS	Mus musculus.		
FH	Key	Location/Qualifiers	
FT	CDS	3..2160	
FT	/*tag= a		
FT	sig_peptide	3..56	
FT	*tag= b		
FT	mat_peptide	57..2156	
FT	*tag= c		
PN	707107979-A.		
PD	25-APR-1995.		
PF	15-AUG-1994; 214227.		
PR	19-AUG-1993; JP-227881.		
PA	(KINO-) KINOSEI PEPTIDE KENKYUSHO KK.		
DR	UPI_95-190179/25.		
P	P-SDS; R90443.		
FN	New DNA encoding an oviduct-specific glycoprotein - useful for		
FS	recombinant protein production in high quantities.		
PT	Claim 1; Pages 11-14; 22pp; Japanese.		
CC	Q090443 encodes R73992 murine oviduct specific glycoprotein (MOGP).		
CC	The cDNA can be used for the commercial recombinant prodn. of		
CO	MOGP in high quantities.		
SO	Sequence 2504 BP; 678 A; 616 C; 538 G; 672 T;		
<hr/>			
Query Match	5.5%; Score 91; DB 15; Length 2504;		
Best Local Similarity	58.5%; Pred. No. 5.83e-42;		

Matches 363; Conservative 0; Mismatches 254; Indels 3; Gaps 2;

Db 60 aaactggtgtgatttaccatattggagacatagtcggccgacctgctccatcaty 119  
 Oy 93 AAACGCGTGTGCTACTTACCAACACGCGCCAGTACAGCAGGGGAGGCTGCTCTCG 152

Db 120 ccccatgacatgacccctcttcttgtaacgcatctgatatcttgcttgctcaatgagc 179  
 Oy 153 CCCAAGGACCTTGGACCCAGCTTGGACCCACCTCATCTAGCCCTTGCTGGCATGACC 212

Db 180 aacaatcagattgttgcgaagaatctgcagatgaaacgttctctatccagatcaac 239  
 Oy 213 AACCAACGAGCT-GAG--CACCACTAGTGAAGAGAGACACTCTTACAGGAGTTCAAT 269

Db 240 aaactcaagaaagaaacagagagctgaaacacgtctgcatcttgaggggtggaacttc 299  
 Oy 270 GGCCTGAAGAAATGAATGCCAAGCTGAAGACCTGTTAGCCATGGAGGCTGGAAATTC 329

Db 300 ggcacatacaggttcaacgcatgtctctcacaacttgcacacgggaaagtattatgac 359  
 Oy 330 AGCACTCAGAAAGTTACAGATATGATGTAGCCACGCGCAACACCGTACAGACTTGTCAAC 389

Db 360 tcaagcacaatcctctctgtaagaatataatgattgtgtgtcttgaccttctctctgtac 419  
 Oy 390 TCGGCAATGAGTTCTTGCGCAAAATACAGCTTTGACGGCTTGACCTTGACTGGAGATAC 449

Db 420 cctgagactacagagagcagtcgccacacagacgggtgaaattctctctcttaataagag 479  
 Oy 450 CCAGGAAGCCAGGAGAGGCTCGCTAGCAAGAGCGCTTACAAACCTGTTACAGAGAC 509

Db 480 ctccagattgctctcagagagagccctgtcactcagaccccaagcggtgctgtctgag 539  
 Oy 510 TTGGCCAATGCTTCCAGAGAGAACCCAGACCTCAGGGAAGAAACGCTTCTGTGAGT 569

Db 540 gctgtgtctctgagatcccaagcatcatcactagctcttaagagcgctctcttaaga 599  
 Oy 570 GCAGGGTTCCAGCTGGCGACACCTATGATGATGCTGATAGAGTGGACAAATCGCC 629

Db 600 aggcgttggattcattatgattgttctcttaagcttaccatggaagtggaaagt 659  
 Oy 630 CAGAACCTGGATTTGTCAACCTTATGGCCTTACGACTTCATGAGCTCTTGGAGAAAGTC 689

Db 660 acaagacacaatagctctct 679  
 Oy 690 ACGGACATTAACAGCCCTCT 709

RESULT 5  
 ID 090442 standard; cDNA to mRNA; 1994 BP.  
 AC 090442;  
 DT 10-JAN-1996 (first entry)  
 DE Bovine oviduct specific glycoprotein cDNA.  
 KW Bovine oviduct specific glycoprotein; recombinant production;  
 KM BOP; ds.  
 OS Bos taurus.  
 FS Bos taurus.  
 FT Key Location/Qualifiers  
 FT CDS 2..1615  
 FT /\*tag- a  
 FT sig-peptide 2..55  
 FT /\*tag- b  
 FT mat-peptide 56..1612  
 FT /\*tag- c  
 FT J07107979-A.  
 PN J07107979-A.  
 PD 25-APR-1995.  
 PR 15-AUG-1994; 214227  
 PR 19-AUG-1993; JP-227881.  
 PA (KINO-) KINOSEI PEPTIDE KENKUSHO KK.  
 DR MPI; 95-190179/25.  
 DR P-PSDB; R73991.  
 PT New DNA encoding an oviduct-specific glycoprotein - useful for  
 PT recombinant protein production in high quantities.  
 PS Claim 1; Pages 7-9; 22pp; Japanese.

CC 090442 encodes R73991 bovine oviduct specific glycoprotein (BOP).  
 CC The cDNA can be used for the commercial recombinant prodn. of  
 CC BOP in high quantities.  
 SO Sequence 1994 BP; 445 A; 536 C; 520 G; 493 T;

Query Match 5.0%; Score 83; DB 15; Length 1994;  
 Best Local Similarity 57.9%; Pred. No. 1.33e-36;  
 Matches 359; Conservative 0; Mismatches 258; Indels 3; Gaps 2;

Db 59 aagctggtgttatttaccatattggatcagtcgcccggcctgctcaatctg 118  
 Oy 93 AAACGCTGTCTACTTCTACCACTGGGCGCCAGTACAGACAGGGGAGGCTCGCTCTCG 152

Db 119 cctcggaatcggacccgcttctctcagccacactggatcttgcttgctgatgagc 178  
 Oy 153 CCCAAGAGCTTGGAGCCACGAGCTTGGACCCACCTCATCTAGCTTGGCTGGATGACC 212

Db 179 aacaatcagattgttcttaagatcccaagatgaaatcctctaccacagatcaac 238  
 Oy 213 AACCAACGAGCTGAGCACACACTGATGGAATGACAGAGC--TC-TCTACAGAGATTCAAT 269

Db 239 aagctcaagagagaaagagaggtctgaaacgctgtctcaccatcgagggttgaaactt 298  
 Oy 270 GGCCTGAAGAAATGAATCCCAAGCTGAAGACCTGTTAGCCATCGGAGCTGGAATTC 329

Db 299 ggcacagtgaggttcaacacagatgtctcagcgttttccaacgggaaaggtctgcaat 358  
 Oy 330 AGCACTCAGAAATGATACAGATATGATGTAGCCACGCGCAACACCTGATGTTCAAC 389

Db 359 tcaagtatgcctctcagagacacatgagcttgatgtgtctgagccttctctctgtac 418  
 Oy 390 TCGGCAATGAGTTCTTCCGCAAAATACAGCTTTGACGGCTTGACTGATGGAGATAC 449

Db 419 cctgagactcaagagagccctgctgagcagctgagacacttctctcttaactgaagag 478  
 Oy 450 CCAGGAAGCCAGGAGAGGCTCGCTAGCAAGAGAGCGCTTACAAACCTGTTACAGAGAC 509

Db 479 ctccgagagccttcaagaatgagccagatcccatgctcgaagcgtctgctgtct 538  
 Oy 510 TTGGCCAATGCTTCCAGAGAGAACCCAGACCTCAGGGAAGAAACGCTTCTGTGAGT 569

Db 539 gctgcgtctctgagagaccccatgtctgccaagaagctgaagaagcagcctctgagc 598  
 Oy 570 GCAGGGTTCCAGCTGGCGACACCTATGATGATGCTGATAGAGTGGACAAATCGCC 629

Db 599 agactccgtgattcattatgattgttctcttaagcttaccatggaagtggaaagt 658  
 Oy 630 CAGAACCTGGATTTGTCAACCTTATGGCCTTACGACTTCATGAGCTCTTGGAGAAAGTC 689

Db 659 acaagacacaatagccctct 678  
 Oy 690 ACGGACATTAACAGCCCTCT 709

RESULT 6  
 ID 010572 standard; DNA; 1047 BP.  
 AC 010572;  
 DT 09-APR-1991 (first entry)  
 DE Human Natriuretic Peptide Receptor B.  
 KW NPB; ANP; BNP; CNP; Kidney failure; heart failure; protein kinase;  
 KW hyperaldosteronism; glaucoma; guanyl cyclase.  
 OS Homo sapiens.  
 FS Homo sapiens.  
 FT Key Location/Qualifiers  
 FT Peptide 1..22  
 FT /label- signal sequence  
 FT Protein 12  
 FT /label- mature NPBR  
 FT Domain 23..455  
 FT /label- extracellular domain  
 FT /note- "binds natriuretic peptides A, B and C"  
 FT Domain 456..456  
 FT /label- transmembrane domain  
 FT Domain 479..1047





OY 93 AACGTGCTGCTACTTCACCACTGGGCCAGATACAGAGGGGAGGCGCTTCTG 152  
DB 120 ccccgtagctgagatcccttctgttacaacatgatatgttgccttgcctgtaggc 179  
OY 133 CCCAAGAGACTTGAGACCCCGCTTTGGACCCACTCATCTACGCTCTGCTGGCATGACC 212  
DB 180 aacaatcagatgtgtgcacataatctccagagatgagaataatctctatccagatc 239  
OY 213 AACCAACAGACTGACGACCACTGAG-TGA-ATGACGAGACTCTCTACAGAGATTCAAT 269  
DB 240 aacatcagagagagagacagagccttgaataactactgtctgttggaggtcgaactc 299  
OY 270 GGCCTGAAGAAATGATCCAGACCTGTAAGACCTGTTCATCGAGAGCTGGAATTTC 329  
DB 300 ggcacatcaggttcacacatgctgtctgcaccttgcagcgcgtgaaatatttgc 359  
OY 330 AGCACTCAGAAATTCACAGATATGGTAGCCAGCCACACACCGTCAGCTTTGTCAAC 389  
DB 360 tcaagtgtatcctctctggaacacatggttgaatggtgcttgcctctctctgtac 419  
OY 390 TCGGCATCAGGTCTTCTGGCAATATACAGCTTGAAGCGCTTGACCTTGACGTGGAGTAC 449  
DB 420 cctggaactcagagcagccc 439  
OY 450 CCAGGAAGCCAGGGAGGCC 469

## RESULT 8

ID N81164 standard; DNA: 204 BP.  
AC N81164.  
DT 08-NOV-1990 (first entry)  
DE Base substituted E.coli beta-galactosidase alpha-fragment.  
KW E.coli beta galactosidase alpha-fragment; base substitutions; ss.  
OS Escherichia coli.  
FH Key Location/Qualifiers  
FT misc-feature 19..69  
FT /tag- a  
FT /function-multiple cloning site  
FT primer\_bind 187..204  
FT /tag- b  
PN EP-285123-A.  
PD 05-MAY-1988.  
PR 30-MAR-1988; 105163.  
PR 03-APR-1987; US-034819.  
PA (SUSO) SUDOM SOKERI OY.  
PI Lehtovaara P, Knowles J, Kolvola A, Bamford J, Reinkainen T.  
DR WPI: 88-279927/40.  
PT Introducing random point mutations into nucleic acids -  
PT by prepn of single stranded template, annealing a primer, elongation,  
PT misincorporation, completion of molecules and screening.  
PS Disclosure; P: English.  
CC Random point mutations were introduced into the alpha fragment of  
CC E.coli beta-galactosidase. The wild type sequence was obtained as a  
CC single stranded template and an oligonucleotide was hybridized to  
CC it to generate a popn of DNA molecules which terminate at all  
CC possible nucleotide positions within a specified region. The  
CC variable 3' ends generated in this way are used as primers for  
CC reverse transcriptase. Nucleotides are misincorporated by the  
CC transcriptase and the molecules are completed to forms that can be  
CC amplified and then expressed in a suitable host-vector system.  
CC The sequence covers all 176 diff base substitutions, most of which  
CC occurred singularly in any given mutant.  
CC See also P80575.  
SQ Sequence 204 BP; 21 A; 47 C; 17 G; 11 T; 108 Others;

Query Match 3.1%; Score 51; DB 1; Length 204;  
Best Local Similarity 15.4%; Pred. No. 5,90e-16;

Matches 18; Conservative 62; Mismatches 36; Indels 1; Gaps 1;

DB 75 yydcdhycgcgymrtthhirmbhvyrdynsdaaayccyrrsvkxkdcynachdd 134  
OY 628 CCCAGAACTGGATTGTGCAACTTATGCGCTACAGACTTCATGCGCTTGGAGAAAG 687

DB 135 hyvbbvynvhnncnccbnhvcnvhbnhrnwayvrhdarriddhcevhc 191  
OY 668 TCACGGGCAATACAGCCCCCTC-TACAGAGGCAAGAGAGATGCTGACAGACC 743

## RESULT 9

ID Q51746 standard; cDNA: 91 BP.  
AC Q51746.  
DT 31-MAY-1994 (first entry)  
DE Oligonucleotide probe MK14-A  
KW Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;  
KW ss.  
OS Synthetic.  
PN EP-571911-A.  
PD 01-DEC-1993.  
PR 24-MAY-1993; 108325.  
PR 26-MAY-1992; US-889651.  
PA (BECT ) BECTON DICKINSON CO.  
PI Shank DD, Spears PA.  
DR WPI: 93-378844/48.  
PT New oligo:nucleotide probes specific for Mycobacteria - used for  
PT detection and amplification of Mycobacteria nucleic acid in  
PT samples  
PS Claim 3; Page 14; 23pp; English.  
CC Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14  
CC (Q51735). It hybridized to all spp. of mycobacteria tested, but  
CC cross reacted to a few non-mycobacterial spp. The probe may  
CC be useful as an initial screen for mycobacterial infection.  
CC See also Q51735-45 and Q51747-59.  
SQ Sequence 91 BP; 5 A; 17 C; 15 G; 4 T;

Query Match 2.7%; Score 45; DB 9; Length 91;  
Best Local Similarity 2.0%; Pred. No. 2,64e-12;

Matches 1; Conservative 47; Mismatches 3; Indels 0; Gaps 0;

DB 10 gsvhsyyvvvhhshvhhvhhvsvvvvhhvhhvhhvhhvhyvsv 60  
CP 1386 GGCCTGTGGCGAGCTTGTGTGGAACAGCCGCCCGCTGCACAGCTGAGA 1336

## RESULT 10

ID Q51746 standard; cDNA: 91 BP.  
AC Q51746.  
DT 31-MAY-1994 (first entry)  
DE Oligonucleotide probe MK14-A  
KW Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;  
KW ss.  
OS Synthetic.  
PN EP-571911-A.  
PD 01-DEC-1993.  
PR 24-MAY-1993; 108325.  
PR 26-MAY-1992; US-889651.  
PA (BECT ) BECTON DICKINSON CO.  
PI Shank DD, Spears PA.  
DR WPI: 93-378844/48.  
PT New oligo:nucleotide probes specific for Mycobacteria - used for  
PT detection and amplification of Mycobacteria nucleic acid in  
PT samples  
PS Claim 3; Page 14; 23pp; English.  
CC Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14  
CC (Q51735). It hybridized to all spp. of mycobacteria tested, but  
CC cross reacted to a few non-mycobacterial spp. The probe may  
CC be useful as an initial screen for mycobacterial infection.  
CC See also Q51735-45 and Q51747-59.  
SQ Sequence 91 BP; 5 A; 17 C; 15 G; 4 T;

Query Match 2.5%; Score 42; DB 9; Length 91;  
Best Local Similarity 3.8%; Pred. No. 1,60e-10;

Matches 2; Conservative 45; Mismatches 5; Indels 0; Gaps 0;

DB 12 svhsyyvvvhhshvhhvhhvsvvvvhhvhhvhhvhhvhyvsv 63  
OY 149 COTGCCAAGACTTGACCCAGCCTTTCACCCACACTCATATACGCTTC 200

RESULT 11  
ID Q70468 standard; DNA; 114 BP.  
AC Q70468;  
DT 05-APR-1995 (first entry)  
DE Generic DNA sequence to generate a random TSAR peptide library.  
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;  
KW effector domain; concatenated heterofunctional protein; linker;  
KW direct; rapid; detection; screening; treatment; generic; ss.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT misc\_feature 55..60  
FT /tag- a  
FT /note- "this sequence represents 'Z'; Z can be a  
FT sequence of 6, 9 or 12 nucleotides (see  
FT comments)"  
PN WO9418318-A.  
PD 18-AUG-1994.  
PF 01-FEB-1994; U00977.  
PR 01-FEB-1993; US-013416.  
PR 30-DEC-1993; US-176500.  
PR 31-JAN-1994; US-189331.  
PA (UYN-) UNIV NORTH CAROLINA.  
PI Fowlkes DM, Kay BK.  
PI WPI; 94-279739/34.  
DR P-PSDB; R65154.  
PT Identifying proteins or peptide(s) which bind a ligand - by  
PT screening a recombinant vector library expressing fusion proteins  
PS comprising a binding domain and an effector domain  
PS Disclosure; Page 35; 25pp; English.  
CC Q70468 is a generic DNA sequence used to generate random TSAR (Totally  
CC Synthetic Affinity Reagents) peptides. This generic formula can also be  
CC represented as follows: X(NNB)11(TGC)(NNB)6Z(NNB)7(TGC)(NNB)10Y. X  
CC and Y are flanking restriction sites (X is not the same as Y) that are  
CC not specified further. Other generic sequences are shown in Q70466-68.  
CC Other specific peptides generated by these generic sequences are shown in  
CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,  
CC comprising at least two functional regions - a binding domain with  
CC affinity for a ligand and a second effector peptide portion that is  
CC chemically or biologically active. They may further comprise a linker  
CC peptide between the 2 domains. The oligonucleotides are also designed so  
CC that the expressed peptide contains 2 or 4 cysteine residues positioned  
CC in, or flanking, the unpredicted or variant residues. These residues  
CC confer some degree of conformational rigidity to the peptides. The TSARs  
CC deliver a chemically or biologically active moiety, eg. metal ion,  
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the  
CC cell. They can also replace the function of macromolecules, eg.  
CC monoclonal or polyclonal antibodies and therefore circumvent the need  
CC for complex methods of hybridoma formation or in vivo antibody  
CC production. The TSARs are easily characterised and have designed activity  
CC allowing direct and rapid detection in a screening process.  
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;  
Query Match 2.3%; Score 38; DB 12; Length 114;  
Best Local Similarity 4.5%; Pred. No. 3.39e-08;  
Matches 5; Conservative 34; Mismatches 73; Indels 0; Gaps 0;

KW E.coli beta galactosidase alpha-fragment; base substitutions; ss.  
OS Escherichia coli.  
FH Key Location/Qualifiers  
FT misc\_feature 19..69  
FT /tag- a  
FT /function-multiple cloning site  
FT primer\_bind 187..204  
FT /tag- b  
PN EP-285123-A.  
PD 05-MAY-1988.  
PF 30-MAR-1988; 105163.  
PR 03-APR-1987; US-034819.  
PA (SUSO) SNOMEN SOKERI OY.  
PI Lehtovaara P, Knowles J, Kolvula A, Bamford J, Reinikainen T;  
PI WPI; 88-279927/40.  
PT Introducing random point mutations into nucleic acids -  
PT by prep of single stranded template, annealing a primer, elongation,  
PT misincorporation, completion of molecules and screening.  
PS Disclosure; P; English.  
CC Random point mutations were introduced into the alpha fragment of  
CC E.coli beta-galactosidase. The wild type sequence was obtained as a  
CC single stranded template and an oligonucleotide was hybridised to  
CC it to generate a popn of DNA molecules which terminate at all  
CC possible nucleotide positions within a specified region. The  
CC variable 3' ends generated in this way are used as primers for  
CC reverse transcriptase. Nucleotides are misincorporated by the  
CC transcriptase and the molecules are completed to forms that can be  
CC amplified and then expressed in a suitable host-vector system.  
CC The sequence covers all 176 diff base substitutions, most of which  
CC occurred singularly in any given mutant.  
SQ See also P80575.  
CC Sequence 204 BP; 21 A; 47 C; 17 G; 11 T; 108 Others;  
Query Match 2.3%; Score 38; DB 1; Length 204;  
Best Local Similarity 11.5%; Pred. No. 3.39e-08;  
Matches 11; Conservative 49; Mismatches 36; Indels 0; Gaps 0;

Db 91 thyrmbvnyridnrsdaawycyrysvkydcynachdhhyvbbvynvnhnn 150  
Cp 1150 TATGCGCCCTGTGACGAGAAAGCGCAAGTCATCTAGTCCAGCCAGACCATG 1091  
Cp 1090 GCCCGGCCAGTCCCTTCTGCTGATGATGCTGACC 1055

RESULT 13  
ID Q70467 standard; DNA; 114 BP.  
AC Q70467;  
DT 05-APR-1995 (first entry)  
DE Generic DNA sequence to generate a random TSAR peptide library.  
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;  
KW effector domain; concatenated heterofunctional protein; linker;  
KW direct; rapid; detection; screening; treatment; generic; ss.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT misc\_feature 55..60  
FT /tag- a  
FT /note- "this sequence represents 'Z'; Z can be a  
FT sequence of 6, 9 or 12 nucleotides (see  
FT comments)"  
PN WO9418318-A.  
PD 18-AUG-1994.  
PF 01-FEB-1994; U00977.  
PR 01-FEB-1993; US-013416.  
PR 30-DEC-1993; US-176500.  
PR 31-JAN-1994; US-189331.  
PA (UYN-) UNIV NORTH CAROLINA.  
PI Fowlkes DM, Kay BK.  
PI WPI; 94-279739/34.  
DR P-PSDB; R65153.  
PT Identifying proteins or peptide(s) which bind a ligand - by  
PT screening a recombinant vector library expressing fusion proteins









CC In, or flanking, the unpredicated or variant residues. These residues  
CC confer some degree of conformational rigidity to the peptides. The TSARS  
CC or compuns, comprising a TSAR binding domain can be used *in vivo* to  
CC deliver a chemically or biologically active moiety, eg. metal ion,  
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the  
CC cell. They can also replace the function of macromolecules, eg.  
CC monoclonal or polyclonal antibodies and therefore circumvent the need  
CC for complex methods of hybridoma formation or *in vivo* antibody  
CC production. The TSARS are easily characterised and have designed  
CC activity allowing direct and rapid detection in a screening process.  
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match	2.28;	Score 36;	DB 12;	Length 114;
Best Local Similarity	3.68;	Pred. NO. 4.64e-07;		
Matches	4;	Conservative	34;	Mismatches 74;
			Indels	0;
			Gaps	0;

Db	3	bnnbnnbnnbnnb	tcgcnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnn	62
Oy	155	CAAGACCTTGGACCCCAACCCTTGGACCACTCATCTAGCCCTTGGCATGACCAA	214	
Db	63	bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnb	tcgcnbnnbnnb	114
Oy	215	CCACCAGCTGACACCACTGATGGAAATGACGAACACTCTCTACCAAGAGATTC	266	

Search completed: Sun Jun 29 20:48:37 1997  
Job time : 185 secs.



Release 2.1d John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K  
Distribution rights by IntelliGenetics, Inc.

MPsrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm

```
Run on: Sun Jun 29 20:49:03 1997; MasPar time 1172.65 Seconds
```

Tabular output not generated.

```

Title: >US-08-663-618A-3
Description: (1-1656) from US08663618A.seq
Perfect Score: 1656
N.A. Sequence:
Comp: 1 GCTGCAAGCTCTCCGCTGAGC.....TGTGCCCCCTCAAAAAAAAAA 1656
CGAGCTCGAGCGGCGACATCG.....ACACGGGAGATTTTTTTTT

```

Scoring table: TABLE default

Nmatch STD : Dbase 0; Query 0

Searched: 886179 seqs, 320143548 bases x 2

```
Post-processing: Minimum Match 0%
                  Listing first 45 summaries
```

Database:

Database:

EST1-STS

1. EST1 2. EST2 3. EST3 4. EST4 5. EST5 6. EST6 7. EST7 8. EST8 9. EST9 10. EST10 11. EST11 12. EST12 13. EST13 14. EST14 15. EST15 16. EST16 17. EST17 18. EST18 19. EST19 20. EST20 21. EST21 22. EST22 23. EST23 24. EST24 25. EST25 26. EST26 27. EST27 28. EST28 29. EST29 30. EST30 31. EST31 32. EST32 33. EST33 34. EST34 35. EST35 36. EST36 37. EST37 38. EST38 39. EST39 40. EST40 41. EST41 42. EST42 43. EST43 44. EST44 45. EST45 46. EST46 47. EST47 48. EST48 49. EST49 50. EST50 51. EST51 52. EST52 53. EST53 54. EST54 55. EST55 56. EST56 57. EST57 58. EST58 59. EST59 60. EST60 61. EST61 62. EST62 63. EST63 64. EST64 65. EST65 66. EST66 67. EST67 68. EST68 69. EST69 70. EST70 71. EST71 72. EST72 73. EST73 74. EST74 75. EST75 76. EST76 77. EST77 78. EST78 79. EST79 80. EST80 81. EST81 82. EST82 83. EST83 84. EST84 85. EST85 86. EST86 87. EST87 88. EST88 89. EST89 90. EST90 91. EST91 92. EST92 93. EST93 94. EST94 95. EST95 96. EST96 97. EST97 98. EST98 99. EST99

EST1-STS-TWO

Database:

Result No.	Score	Match	Query Length	DB	ID	Description	Pred. No.
SUMMARIES							
c 1	345	20.8	428	151	T94579	yesjg05.s1 Homo sapie	0.00e+000
c 2	245	14.8	405	151	T94272	yesjg05.r1 Homo sapie	0.00e+000
c 3	215	13.0	241	250	HS6114	human STS SHGC-11364.	0.00e+000
c 4	215	13.0	241	202	G29124	human STS SHGC-11364.	0.00e+000
c 5	79	4.8	405	41	H43201	yp05a1.r1 Homo sapie	1.44e-100
c 6	78	4.7	541	113	N75653	yr28c01.r1 Homo sapie	1.34e-98
c 7	78	4.7	541	207	N75653	yr28c01.r1 Homo sapie	1.34e-98
c 8	73	4.4	533	144	G24826	ycl2b03.s1 Homo sapie	7.80e-89
c 9	73	4.4	533	200	G44226	human STS WI-14165.	7.80e-89
c 10	71	4.3	311	79	N40107	ym72b09.r1 Homo sapie	5.83e-85
c 11	71	4.3	336	169	W51302	ym72b09.r1 Homo sapie	5.83e-85
c 12	66	4.0	266	35	H24821	yl42c01.r1 Homo sapie	2.27e-75
c 13	65	3.9	396	164	R30530	yl79c12.r1 Homo sapie	1.81e-73
c 14	65	3.6	306	215	W10705	ma47g07.r1 Soares mou	4.65e-64
c 15	60	3.6	306	174	W10705	ma47g07.r1 Soares mou	4.65e-64
c 16	60	3.6	421	139	R74158	yl99g05.r1 Homo sapie	4.65e-64
c 17	60	3.6	438	76	N26985	ym73b09.s1 Homo sapie	4.65e-64
c 18	60	3.6	586	170	W53184	md13a03.r1 Soares mou	4.65e-64
c 19	59	3.6	589	193	W97818	mg03c04.r1 Soares mou	3.38e-62
c 20	56	3.4	320	144	T66009	ycl2b03.r1 Homo sapie	1.17e-56
c 21	291	3.0	299	151	T91693	ye25b05.r1 Homo sapie	4.86e-44
c 22	43	2.6	459	31	H10989	ym07a06.s1 Homo sapie	1.22e-33
c 23	41	2.5	461	56	H94102	ym59g11.r1 Homo sapie	2.80e-30
c 24	38	2.3	392	95	N64805	yz31c01.s1 Homo sapie	2.36e-25
c 25	36	2.2	335	47	CE1K104H2F	C.elegans CDNA clone	3.71e-22
c 26	35	2.1	280	98	R03212	Pk05a03.s1 Caenorhabd	1.38e-20
c 27	35	2.1	360	46	CE1K084F3F	C.elegans CDNA clone	1.38e-20
c 28	34	2.1	360	47	CE1K104C7F	C.elegans CDNA clone	4.86e-19
c 29	34	2.1	360	45	CE1K059D7F	C.elegans CDNA clone	4.86e-19
c 30	34	2.1	429	45	CE1K009B7F	C.elegans CDNA clone	4.86e-19
c 31	33	2.0	360	45	CE1K066G3F	C.elegans CDNA clone	1.63e-17
c 32	33	2.0	360	47	CE1K109D2F	C.elegans CDNA clone	1.63e-17
c 33	32	1.9	323	16	AA067770	mm24c01.r1 Stratagene	5.19e-16
c 34	32	1.9	323	244	NMA667770	mm24c01.r1 Stratagene	5.19e-16
c 35	31	1.9	379	224	CEC4471	C.elegans CDNA clone	1.56e-14
c 36	31	1.9	379	39	C08467	C.elegans CDNA clone	1.56e-14
c 37	28	1.7	280	231	HSA62630	zm02h01.r1 Stratagene	2.89e-10
c 38	28	1.7	280	14	AA062630	zm02h01.r1 Stratagene	2.89e-10
c 39	26	1.6	331	21	AA054927	SMFECBA00SK Brugia ma	1.40e-07
c 40	26	1.6	331	223	BMA54927	SMFECBA00SK Brugia ma	1.40e-07
c 41	25	1.5	447	109	G23262	human STS WI-14048.	2.77e-06
c 42	25	1.5	447	109	R40271	yf80f04.r1 Homo sapie	2.77e-06
c 43	24	1.4	375	185	W18653	mb99b05.r1 Soares mou	4.82e-05

44 24 1.4 375 218 W18653 mb98b05.r1 Soares mou 4.82e-05  
45 24 1.4 450 216 W13185 ma81d02.r1 Soares mou 4.82e-05

ALIGNMENTS

RESULT 1  
LOCUS T94579 428 bp mRNA EST 24-MAR-1995  
DEFINITION Ye31905.s1 Homo sapiens cDNA clone 119384 3'.  
ACCESSION T94579  
NID 9728067  
KEYWORDS  
SOURCE human clone-119384 library-Stratagene lung (#937210)  
vector-pBluescript SK- host-SOLR cells (kanamycin resistant)  
primer-21m3 Rsite1-EcoRI Rsite2-XhoI Normal lung tissue from a 72  
year old male. Cloned unidirectionally. Primer: Oligo dT. Average  
insert size: 1.0 kb; Uni-ZAP XR Vector; 5' adaptor sequence:  
5'-GAATTCGGCAGAG-3'; 3' adaptor sequence:  
5'-CTCGAGTTTATTTTATTTT-3'.

ORGANISM  
Homo sapiens  
Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;  
Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE  
1 (bases 1 to 428)  
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,  
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,  
Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevasakis,E.,  
Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.  
WashU-Merck EST Project  
Unpublished (1995)

TITLE  
JOURNAL  
COMMENT

CONTACT: Wilson RK  
WashU-Merck EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu  
High quality sequence stops: 241  
Source: IMAGE Consortium, LLNL  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES  
source  
1..428  
/organism="Homo sapiens"  
/clone="119384"

BASE COUNT 112 a 103 c 124 g 76 t 13 others

ORIGIN  
Query Match 20.8%; Score 345; DB 151; Length 428;  
Best Local Similarity 94.3%; Pred. No. 0.00e+00;  
Matches 383; Conservative 0; Mismatches 18; Indels 5; Gaps 5;

Db 7 gangggcacaacaaagatttttttgcagtgaaagggcgcc-aggagaccagaa 65  
CP 1646 GAGGGGCACAAACAAAGATTTTTCAGAGGAAAGGGGCGACCCAGAGACCCAGAA 1587  
Db 66 aaaaggaaggcaaggtgtgagcagagaagcgtgataaaggaagacacagaaagccg 125  
CP 1586 AAAAGGAAGGCAAGGCTGAAGCAAGAAAGCTGTGATPAAGAGACACAGAAAGGCGTG 1527  
Db 126 caggaagcagatgtcgccccccagggaaacacagagagcgtctagatgatccg 185  
CP 1526 CAGGACCCAGATTCGGCCCCCGGAGGAAACCCAGAGAGGAGCTGTAGAGTATCTCG 1467  
Db 186 gccacgctcaaaagcttgagcttgagagggctttaa-gactcaattccaggtgacatt 244  
CP 1466 GCCCAGCTCAAAAGCTGGGAGCTGAGAGGCGCTTACGAGCTCAATTCAGATTCAGCATTT 1407  
Db 245 gaagagntgtctgaacacagcgctntcgcgagcttt-ctgnaacacacmgccccgntgc 303  
CP 1406 GCGAGGATTCCTGAACCAAGGCGCTTGTGGGAGCTTTTGTGGAAACAGCCGCCCGCTGC 1347  
Db 304 acaactgtagaagctgacgcgttcccgagagnttgagatagagccatcaagctt-ccctg 362

CP 1346 ACAGCGTAGAAGCGTAGACCGTTCGCCAGAGGATGGGATGAGACCATACAGCTTTCCTCG 1287  
Db 363 gcagaaagctttctgtncaggggctngagcagtcagctcagctcaga 408  
CP 1286 GCAGACGTGTCTGTCCAGGGCTGGGGCCATGCT-CAGGTTGAGA 1242

ALIGNMENTS

RESULT 2  
LOCUS T94272 405 bp mRNA EST 24-MAR-1995  
DEFINITION Ye31905.r1 Homo sapiens cDNA clone 119384 5'.  
ACCESSION T94272  
NID 9727760  
KEYWORDS  
SOURCE human clone-119384 library-Stratagene lung (#937210)  
vector-pBluescript SK- host-SOLR cells (kanamycin resistant)  
primer-M13RPI Rsite1-EcoRI Rsite2-XhoI Normal lung tissue from a 72  
year old male. Cloned unidirectionally. Primer: Oligo dT. Average  
insert size: 1.0 kb; Uni-ZAP XR Vector; 5' adaptor sequence:  
5'-GAATTCGGCAGAG-3'; 3' adaptor sequence:  
5'-CTCGAGTTTATTTTATTTT-3'.

ORGANISM  
Homo sapiens  
Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;  
Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE  
1 (bases 1 to 405)  
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,  
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,  
Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevasakis,E.,  
Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.  
WashU-Merck EST Project  
Unpublished (1995)

TITLE  
JOURNAL  
COMMENT

CONTACT: Wilson RK  
WashU-Merck EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu  
High quality sequence stops: 214  
Source: IMAGE Consortium, LLNL  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES  
source  
1..405  
/organism="Homo sapiens"  
/clone="119384"

BASE COUNT 64 a 113 c 102 g 115 t 11 others

ORIGIN  
Query Match 14.8%; Score 245; DB 151; Length 405;  
Best Local Similarity 91.8%; Pred. No. 0.00e+00;  
Matches 371; Conservative 0; Mismatches 14; Indels 20; Gaps 20;

Db 1 tctgaacctgagcatgcccagccctggacagaagcgttgcagggcagaagctgat 60  
QY 1242 TCTGAACCTGAGCATGCCCCGACCCCTGACAAAGACACTTTCGCAAGGCGCAAAAGCTGAT 1301  
Db 61 gggcttataccaactcctcgggaagcgttcagcttctaacagctgtgcagcggggcg-ctg 119  
QY 1302 GGGCTTATCCCAATCTCTGGGAAGCGTCCAGCTTCTACAGCTGTGACAGCGGGGGGCTG 1361  
Db 120 ttccagcaagctgtccggcag-cctgtgtctcagcaactcctgtcaaatgtcgacactg 178  
QY 1362 TTCACAGCAAGGTGCCCGACAGCGCTGTGTTCAGCAACTCTGCAAAAGCTGCACACTGG 1421  
Db 179 gaattgaagtcgctaaagccccccagcttccagctttagagctggggccagggatcaatc 238  
QY 1422 -AATTAGTCGCTAAAGGCCCTTCACATCCCACTTTGAGCTGGGCGCCAGG-AT-CATCTC 1478  
Db 239 ttcaagccttgccttctcctgttttcccttggggggcgcaactctgtgnttcccttgcagcg 298  
QY 1479 T-ACAGCCT-GCCT-CCTGGG-TTTTCCCTGGGGG-CCGCAATCTGG-CTCCT-GCAGGC 1531



Df	299	ccttttggnttccttccatnccaaggtttnttctcgaagtntgcttcctt	358
Oy	1532	TTTTT- TTTGCTT- TTTTT- TTTTT- TTTTT- TTTTT- TTTTT-	1566
Df	359	tnttggttccctcgaggtcccttcattngcaaanfaat	403
Oy	1587	TTTTT- TTTTT- TTTTT- TTTTT- TTTTT- TTTTT-	1628
RESULT	3	standard; DNA; STS; 241 BP.	
ID	HSG124		
AC	G29124:		
NI	q1592433		
Df	04-OCT-1996	(Rel. 49, Created)	
Df	05-OCT-1996	(Rel. 49, Last updated, Version 2)	
Df	human STS SHGC-17364.		
KW	primer: sequence tagged site; STS sequence.		
OS	Homo sapiens (human)		
OC	Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;		
CC	Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.		
RN	[1]		
RP	1-241		
RA	Myers R.M.;		
RT	;		
RL	Unpublished.		
CC	Contact: Richard M. Myers, Stanford Human Genome Center (SHGC)		
CC	Stanford University School of Medicine Department of Genetics,		
CC	M-444, Stanford, CA 94305, USA Tel: 4157259687 Fax: 4157259689		
CC	Email: myers@shgc.stanford.edu Primer A: AAAGCCTGATTAAGAAGACC		
CC	Primer B: GCGCACCTGGATTGAGTC STS size: 150 PCR Profile: Initial		
CC	incubation: 94 degrees C for 90 seconds Denaturation: 94 degrees C		
CC	for 15 seconds Annealing: 62 degrees C for 23 seconds		
CC	Polymerization: 72 degrees C for 30 seconds PCR cycles: 30 Thermal		
CC	Cycler: Perkin Elmer 9600 Protocol: Template: 25 ng Primer: each 1		
CC	ul dNTPs: each 200 uM Taq Polymerase: 0.05 units/uL Total Vol: 10		
CC	ul Buffer: MgCl2: 2.5 mM KCl: 50 mM Tris-HCl: 20 mM pH: 8.3		
CC	Prepared with primer pairs derived from T94579--Merck/UniEST.		
CH	Key		
FH	Location/Qualifiers		
FT	source	1..241	
FT	/organism="Homo sapiens"		
FT	/map="1"		
FT	STS	91..240	
FT	primer_bind	91..112	
FT	primer_bind	complement(221..240)	
SQ	Sequence 241 BP: 78 A; 53 C; 77 G; 30 T; 3 other;		
Query Match	13.0%; Score 215; DB 250; Length 241;		
Best Local Similarity	97.0%; Pred. No. 0.00e+00;		
Matches	230; Conservative 0; Mismatches 5; Indels 2; Gaps 2		
Df	7	gangggacaacaaagaattattctgcaagtgaaggagcac- agagaccagaa	65
Cp	1646	GAGGGGCAACAACCAAAATTATTTCAGGTAAAGAGGCGACCCAGAGACACAGAA	1587
Df	66	aaaagaaggaaggaagcttagagcagaagaagcctgatataagaagaacacagaagaagcgt	125
Cp	1586	AAAAGGAAGGCAAGGCTGAGAGCAGAAAAGCCTGGATTAAAGAAABACACACAGAAAAGCCTG	1527
Df	126	cagagcacaagattggcgccccacagagaaaaaccaagagcagcgtgtagaagtactctg	185
Cp	1526	CAGGAGCCAGATTTGGGGCCCCCAGAGAAACCCAGAGGCGGTGATGATCTGG	1467
Df	186	gcccaagcctcaaatgcttgaccttggaagggccttan-gactaatccaaagtgcagca	241
Cp	1466	GCCAGCCTCAAAGCTGGGACTGGAGCTGGAGGGGCTTTAGCGACTCAATTCCAGAGTGACAGCA	1410
RESULT	4	G29124	241 BP
LOCUS	DEFINITION	human STS SHGC-17364.	DNA
ACCESSION	human STS	G29124	ST S
			04-OCT-1996

NID 91592433  
 KEYWORDS STS sequence; primer; sequence tagged site.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata;  
 Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 241)  
 AUTHORS Myers,R.M.  
 JOURNAL Unpublished (1996)  
 COMMENT  
 Contact: Richard M. Myers  
 Stanford Human Genome Center (SHGC)  
 Stanford University School of Medicine  
 Department of Genetics, M-344, Stanford, CA 94305, USA  
 Tel: 4157259687  
 Tel: 4157259687  
 Fax: 4157259689  
 Email: myerseshgc.stanford.edu  
 Primer A: AAGCGCTGATTAAGAGAGAC  
 Primer B: GGTGACCTGTGAATTGAGTC  
 STS size: 150  
 PCR Profile:  
 Initial incubation: 94 degrees C for 90 seconds  
 Denaturation: 94 degrees C for 15 seconds  
 Annealing: 62 degrees C for 23 seconds  
 Polymerization: 72 degrees C for 30 seconds  
 PCR Cycles: 30  
 Thermal Cycler: Perkin Elmer 9600  
 Protocol:  
 Template: 25 ng  
 Primer: each 1 uM  
 dNTPs: each 200 uM  
 Tag Polymerase: 0.05 units/uL  
 Total Vol: 10 uL  
 Buffer:  
 MgCl2: 2.5 mM  
 KCl: 50 mM  
 Tris-HCl: 20 mM  
 pH: 8.3  
 FEATURES  
 source Prepared with primer pairs derived from T94579--Merck/UniLST.  
 location/Qualifiers  
 1..241  
 /organism="Homo sapiens"  
 /map="1"  
 STS  
 primer\_bind 91..240  
 primer\_bind 91..112  
 BASE COUNT 78 a 53 c 77 g 30 t 3 others  
 ORIGIN  
 Query Match 13.0%; Score 215; DB 202; Length 241;  
 Best Local Similarity 97.0%; Pred. No. 0.00e+00;  
 Matches 230; Conservative 0; Mismatches 5; Indels 2; Gaps 2.  
 Db 7 gangggcacaacccaagattattttggcaagttaaggggcagcc-agaagaccagaa 65  
 Cp 1646 GAGGGGCACCAACCAAAATTATTATTGGCAAGTAAAGGGGCACCCAGAGACCCAGAA 1587  
 Db 66 aaaaggaaggaaggaagcgtagaagcagaagcctgtgataaagaagagacacagaagagcgtg 125  
 Cp 1586 AAAAGGAGGCGCAAGGCTGAGCGCGAANAAGCTGCTTAAGGAGAGACACACAGAAAGGCTTG 1527  
 Db 126 caggaagccagatttcgagcccccaggagaaacccagagagcagcgtgtagaagtgaaccttg 185  
 Cp 1526 CAGGAGCCAGATTGGGGCCCCCAGGAGAAACCCAGAGGAGCGAGCGCTGTAGAGTATCTCTGG 1467  
 Db 186 gccagagcctcaaaagcttggaactttgaaggggctttn-gactcaattccagtgagca 241  
 Cp 1466 GCCAGGCTCAAAAGCTGGAGCTGAGGGGGCTTTAGCGACATCAATTCCAGGTGACAGCA 1410

RESULT	5	LOCUS	HA3270	405 bp	mRNA	EST	31-JUL-1995
DEFINITION			yp05a12.r1 Homo sapiens cDNA clone 186526 5' similar to gb:M80927				
ACCESSION			CARTILAGE GLYCOPROTEIN-39. PRECURSOR (HUMAN);.				
KEYWORDS			HA3270				
NID			9919322				
SOURCE			EST.				
			human clone-186526 library-Soares breast 3mbhst vector-pt7n3d (Pharmacia) with a modified polylinker host-CDH10B (ampicillin resistant) primer-M13Rp1 Rsite1-Not I Rsite2-Eco RI Admitt human. 1st strand cDNA was primed with a Not I - Oligo(dt) primer [5' TGTTACCAATCTGAAAGCGAGCGCGCCCTTTTGTGTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7/3 vector (Pharmacia). Library went through one round of normalization to a Cot = 20. Library constructed by Bento Soares and M.Patima Bonaldo.				
ORGANISM			Homo sapiens				
			Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata; Tetrápoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 405)				
REFERENCE			Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kueba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasaki,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.				
AUTHORS			The WashU-Merck EST Project unpublished (1995)				
TITLE							
JOURNAL							
COMMENT							

```

FEATURES
  source
    1..405
    /organism="Homo sapiens"
    /clone="16526"
  106 a 112 c 95 g 86 t 6 others
BASE COUNT
ORIGIN

```

Query Match	4.8%	Score 79	DB 41	Length 405
Best Local Similarity	63.0%	Pred. No. 1,44e-100		
Matches 196	Conservative	0	Mismatches 114	Indels 1
			Gaps	1
Db	1	gctgtctccagatgcctctgacgcgtctcmtctgaccacatcatctaaagctttgcca	60	
QY	145	gcttctctcccaagacattggacacccaccccttctccacccactctatctgacgttgcctg	204	
Db	61	ataataagaacgatcacctcgcacacctgagatggaatgatagtgtaagcgtctacgcatagc	120	
QY	205	gcattaccacacaccacagctgac	264	
Db	121	tcaacacacatcaagaacaggaaccccccaacctgaagactctctgtctgtggaagatgga	180	
QY	265	tcaatggcgctgaagagatggaatcccaagctgaagacccctgttagccatgggaagctgga	324	
Db	181	ncmttgggtctccaaagattttccaaagataagcctccaaacccagagtcgcgcgactttca	240	
QY	325	attttcagacacctcagaaagtccacagatattgattgagccacgacgacacacacacacacac	384	
Db	241	tcaagtgatgaacgcacattctgcgacacctgacttgatgaggtctcgagacacttgctg	300	
QY	385	tcaactctggccatcatgagttcttgcgcacaaatcagctttggaaggc - cttgacattctggaattgg	443	

Db	301	gtctaccctg	311
QY	444	GAGTACCCAGG	454

RESULT	6					
LOCUS	N75653	541 bp	mRNA	EST	02-APR-1996	
DEFINITION	y296c01.r1 Homo sapiens cDNA clone 244128 5' similar to gp:M8092727					
ACCESSION	CARTILAGE GLYCOPROTEIN-39 PRECURSOR (HUMAN);.					
NTID	N75653					
KEYWORDS	g1238231					
SOURCE	EST.					
	human clone-244128 primer-reverse ET library-Soares fetal liver					

spleen INFLS vector-pT73D (Pharmacia) with a modified polylinker host-DH10B (ampicillin resistant), Raitel-Pac I Raitel2-Eco RI Liver and spleen from a 20 week-post conception male fetus. 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' AATCGGAGAGATTATTTAAAGATCTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

1 (bases 1 to 541)  
Hillier, L., Clark, N., Dubuque, T., Eiliston, R., Hawkins, M.,  
Holtman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,  
Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and  
Wilson, R.  
The WashU-Merck EST Project  
Unpublished (1995)

FEATURES	COMMENT
<p> <a href="mailto:estewatson.wustl.edu">Email: estewatson.wustl.edu</a>  <a href="http://www.imageconsortium.org">Source: IMAGE Consortium, LLNL</a>  This clone is available royalty-free through LLNL ; contact the  IMAGE Consortium (<a href="mailto:info@image.llnl.gov">info@image.llnl.gov</a>) for further information.  <a href="#">location/Qualifiers</a> </p>	<p> Contact: Wilson RK  WashU-Merck EST Project  Washington University School of Medicine  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  Tel: 314 286 1800  Fax: 314 286 1810 </p>

	ORIGIN	BASE COUNT	MRNA	4.7%	Score 78;	DB 113;	Length 541;
		149 a	134 c	117 g	136 t	5	others
		<1...541					
		/organism="Homo sapiens"					
		/clone="244128"					
		Best Local Similarity 66.7%; Pred. No. 1,34e-58;					
		Matches 166; Conservative 0; Mismatches 82; Indels 1; Gaps 1;					
Db	27	ctggcgaggtgtagttgcttctgtcgtcttccagggagatctgctcaaacgtgttg	86				
Qy	44	CTGGCGAGGTTCATGTGGCTCGTGTGATGATGCATCGGGCGCTCTGCTCAAAACGTCG	103				
Db	87	ctactttcccaacggtgtcccaagccggcagaaaccaaagaaa-ttcccccctagaatat	145				
Qy	104	CTACTTCACCAACGTGGGCCCAAGTACCAACAGGGGAGGCTCGCTTCTCGGCCAAGGACTT	163				
Db	146	tgaccctctcatgctctcatcatcatctatctatcttgccagcatgtgaaaaacaagaagt	205				
Qy	134	GGAACCCAGCCTTTTGACACCCACCTCATCTTACGCTTGTGGCTGGCATGACCAACACCAAGCT	223				
Db	206	tatcatcaaggacaagaattgaagtgaatgctctaccagaacatatcaacggtctcaaaccaa	265				

OY 224 GAGCACCACGAGTGAAGAGCAGACTCTCTACGAGGATTCAATGGCCTGAAGAAGAT 283  
 Db 266 gaatcccaa 274  
 OY 284 GAATCCCAA 292  
 RESULT 7  
 LOCUS N75653 541 bp mRNA EST 02-APR-1996  
 DEFINITION yv29c01.r1 Homo sapiens CDNA clone 244128 5' similar to gb:M80927  
 CARRIER GLYCOPROTEIN-39 PRECURSOR (HUMAN)).  
 ACCESSION N75653  
 NID 91238231  
 KEYWORDS EST.  
 SOURCE human clone-244128 primer-reverse ET library-Soares fetal liver  
 spleen INFs vector-p773D (Pharmacia) with a modified polylinker  
 host-DH10B (ampicillin resistant) Rsite1-Pac I Rsite2-Eco RI liver  
 and spleen from a 20 week-post conception male fetus. 1st strand  
 cDNA was primed with a Pac I - oligo(dT) primer [5']  
 AACGGAAGATTAATTAAGATCTTTTCTTTTCTTTT 3', double-stranded  
 cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac  
 I and cloned into the Pac I and Eco RI sites of the modified p773  
 vector. Library went through one round of normalization. Library  
 constructed by Bento Soares and M.Fatima Bonaldo.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Eumetazoa; Bilateria; Coelomata;  
 Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;  
 Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria;  
 Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.  
 1 (bases 1 to 541)  
 Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,  
 Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,  
 Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,  
 Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and  
 Wilson,R.  
 TITLE The WashU-Merck EST Project  
 JOURNAL Unpublished (1995)  
 COMMENT  
 Contact: Wilson RK  
 WashU-Merck EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 High quality sequence stops: 441  
 Source: IMAGE Consortium, LNL  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.lnl.gov) for further information.  
 FEATURES  
 source NCBI gi: 1238231  
 Location/Qualifiers  
 1..541  
 /organism="Homo sapiens"  
 /clone="244128"  
 /note="human"  
 mRNA  
 BASE COUNT 149 a 134 c 117 g 136 t 5 others  
 ORIGIN  
 Query Match 4.7%; Score 78; DB 207; Length 541;  
 Best Local Similarity 66.7%; Pred. No. 1.34e-98;  
 Matches 166; Conservative 0; Mismatches 82; Indels 1; Gaps 1;  
 Db 27 ctggagaggtgtatgtctgtctgtctctcagaggagatcgtctacaactgtgttg 86  
 OY 44 CTGGGAGAGTTTATGTCTGTCTGTGATGATCCATGGGCTGCTGCAGCAAAATGCTG 103  
 Db 87 ctacttacaactgtgtccagagacggcagaaacagaaat-ttaaccttgaagaat 145  
 OY 104 CTACTTCACCACTGGGCCAGTACAGACAGGGGAGGCTGCTCTGCCCAAGACTT 163

Db 146 tgacccttcctatgtctctcatcatcattcattcgcagcatcgaacaaacaggt 205  
 OY 164 GGACCCACACTTTTGCACCCACTCTATCTAGCGCTTTCGTGGCATGACCAACACCACT 223  
 Db 206 tatcatcaagaagaagagtgatgtctctaccagacatcaactcaaacca 265  
 OY 224 GAGCACCAGTGAAGTGAATGAGAGAGCTCTACAGAGGATTCATAGGCTGAAGAAGAT 283  
 Db 266 gaatcccaa 274  
 OY 284 GAATCCCAA 292  
 RESULT 8  
 LOCUS T65854 533 bp mRNA EST 20-FEB-1995  
 DEFINITION yca2b03.s1 Homo sapiens CDNA clone 81581 3' similar to sp:S27879  
 S27879 SECRETORY PROTEIN YM-1 PRECURSOR - .  
 ACCESSION T65854  
 NID 9674899  
 KEYWORDS EST.  
 SOURCE human clone-81581 library-Stratagene lung (#937210)  
 vector-pBluescript SK- host-SOLR cells (kanamycin resistant)  
 primer-23ml3 Rsite1-EcoRI Rsite2-XhoI Normal lung tissue from a 72  
 year old male. Cloned unidirectionally. Primer: Oligo dT. Average  
 insert size: 1.0 kb; Uni-ZAP XR Vector; 5' adaptor sequence:  
 5'-GATTCGGCAGCAG-3'; 3' adaptor sequence:  
 5'-CTCGAGTTTCTTTTCTTTTCTTTT-3'.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;  
 Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 1 (bases 1 to 533)  
 Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,  
 Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,  
 Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevasakis,E.,  
 Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.  
 TITLE WashU-Merck EST Project  
 JOURNAL Unpublished (1995)  
 COMMENT  
 Contact: Wilson RK  
 WashU-Merck EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 High quality sequence stops: 307  
 Source: IMAGE Consortium, LNL  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.lnl.gov) for further information.  
 FEATURES  
 source 1..533  
 /organism="Homo sapiens"  
 /clone="81581"  
 BASE COUNT 140 a 123 c 136 g 128 t 6 others  
 ORIGIN  
 Query Match 4.4%; Score 73; DB 144; Length 533;  
 Best Local Similarity 70.1%; Pred. No. 7.80e-89;  
 Matches 131; Conservative 0; Mismatches 55; Indels 1; Gaps 1;  
 Db 250 tggacaagagggttaag 309  
 Cp 1162 TGGATGAGGGGGATTCGGCCCTGTTGAGAGAGCCGCAAGATCATTAATGCTCACT 1103  
 Db 310 gaccagatcatgtctctccagaggttaaatctttaagaactgaacacctgtctccata 369  
 Cp 1102 GCCAGACCATGGCCCCGCCAGTCCCTCTCTCTCAATACCTAGTCTGTTTGAAG 1043  
 Db 370 ctcttcaatcatatagaccacacacactgtgtcccttgactnctagggagagagagag 429  
 Cp 1042 CTCTCCACATCATCAAGGCCACCACT-GGTTGCCGGAAGAGTATAGGACACCTTCTG 984  
 Db 430 atctctg 436



	Db	168	tgacccttcctcgtatctctcatcattcatcttccgcgaagcatcgaaacaagaagt	247
	Oy	164	GGACCCAGACCCTTTTCACACCACCATCTATCACGCTTGGTGCGATCACAACCAACGAGCT	223
	Db	248	tatcataagacgaagtgtagttagttctcacagacatcaaacagctctnaaaa	304
	Oy	224	GAGCACCACTGATGTGGATGATGACGAGACTCTTACCAAGGATTCAATGCGCTGAAGA	280
	RESULT	11		
	LOCUS		W51302	336 bp mRNA EST 29-MAY-1996
	DEFINITION		ma16f08.r1 life Tech mouse brain Mus musculus cDNA clone 304743	5'
	ACCESSION		W51302	similar to gb:X93035 M.musculus mRNA for BRP39 protein (MOUSE)).
	NID		g1339742	
	KEYWORDS		EST.	
	SOURCE		house mouse.	
	ORGANISM		Mus musculus	
			Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;	
			Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;	
			Murinae; Mus.	
	REFERENCE		1 (bases 1 to 336)	
	AUTHORS		Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucab,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theisinger,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.	
	TITLE		The WashU-HMI Mouse EST Project	
	JOURNAL		Unpublished (1996)	
	COMMENT		Contact: Marra M/Mouse EST Project WashU-HMI Mouse EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@wustl.edu This clone is available royalty-free through LNL ; contact the IMAGE Consortium ( <a href="mailto:Info@image.llnl.gov">Info@image.llnl.gov</a> ) for further information. MGI:201503 Seq primer: ETPRimer High quality sequence stop: 148.	
	FEATURES		location/Qualifiers	
	source		1..336	
			/organism="Mus musculus"	
			/strain="C57BL/6J"	
			/note="Vector: PCMV-SPORT2; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. PCMV-SPORT2 vector."	
			/clone="304743"	
			/clone_lib="Life Tech mouse brain"	
			/dev_stage="adult"	
			/lab_host="DH10B"	
			<1..>336	
	BASE COUNT		83 a 100 c 89 g 64 t	
	ORIGIN			
	Query Match		4.3%; Score 71; DB 169; Length 336;	
	Best Local Similarity		69.2%; Pfd. No. 5,83e-85;	
	Matches 137; Conservativity		0; Mismatches 60; Indels 1; Gaps 1;	
	Db	140	caggcttgccggctccttgatgcgtgccagagctgcctcgtcfagcaagtgctgtact	199
	Oy	49	CAGGTTTCATGSETCTGCTGANTGATCCATGGGGGCTCTGTCGAACAACTGEGTGTACT	108
	Db	200	acaccaactggtgccagtatccggaaggaagcaatgaggaagctgtcccagacgacctgag	259
	Oy	109	TCACCAACAGGGGCCAGTAGACAGACAGGGGAGGAGCTCGCTCCGCCCCAAGAGATTGGAC	168
	Db	260	attccctgtgacccatatcatcttaagctttgcaa-catcagccaacaagaagctcaaga	318
	Oy	169	CCAGGCTTTGCACCCACCTCATCTACGCTTGCTGGCATGACCAACACCACTGAGCA	228

DB	319	catccagatgggaatgacg	336	
QY	229	CCACTGAGTGGAAATGACG	246	
RESULT	12			
LOCUS	H24821	266 bp	mRNA	EST
DEFINITION	Y142f01.r1 Homo sapiens cDNA clone 160921 5' similar to gb:M80927			
ACCESSION	CARTIAGE GLYCOPROTEIN-39 PRECURSOR (HUMAN);.			
NID	H24821			
KEYWORDS	G893720			
SOURCE	EST. human clone-160921 library-Saeres breast 3NBHst vector-p7T73D (Pharmacia) with a modified polylinker host-DH10B (ampicillin resistant) primer-M13R1 Rsite1-Not I Rsite2-Eco RI Adult human, 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTACCACGCAAGTGGGACGGCCGCCCTTTTCTTTTCTTTTCTTTT 3'] double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified p7T73 vector (Pharmacia). Library went through one round of normalization to a Cot = 20. Library constructed by Bento Soares and M.Fatima Bonaldo.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Riklin,L., Rohlfing,T., Soares,M., Tan,F., Tveaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.			
TITLE	The Washu-Merck EST Project			
JOURNAL	Unpublished (1995)			
COMMENT	Contact: Wilson RK Washu-Merck EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@wustl.edu High quality sequence stops: 84 Source: IMAGE Consortium, LNL. This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Location/Qualifiers 1..266 /organism="Homo sapiens" /clone="160921"			
FEATURES	source			
BASE COUNT	58 a	80 c	70 g	51 t
ORIGIN	others			
Query Match	4.0%;	Score 66;	DB 35;	Length 266;
Best Local Similarity	68.3%;	Pred. NO. 2.27e-75;		
Matches	127;	Conservative	0;	Mismatches 58; Indels 1; Gaps 1;
Db	14	ggacacagtcacatagaatcctcctggcagcagcgtccctatgacacacaggaacacagt	73	
QY	958	GGGCGCACCCAAACAGAGATCCAGGATCCAGAGCGTCCCTCATCTTCGCGACAACAGT	1017	
Db	74	gggtggtgataagacacacaggaagcgtcaaaagcaagtgacgtactgaaggaacaggc	133	
QY	1018	GGGtGGGCTTTGATGATGTGGAGAGCTTCAAAACCAAGTACGCTATCTGAAGCAAGG	1077	
Db	134	agctgg-cgggcgacatggtatgggccttcnagcctcgatnacttcaagggcctcttcg	192	
QY	1078	GACTGGCGGGGGCCAGTGTGCGGCACGTGACCTTAGATGACTTTGGCGGCTTCTCTGCA	1137	
Db	193	ggcagg	198	
QY	1138	ACCAAG	1143	

LOCUS	13	R55530	396 bp	mRNA	EST	22-MAY-1995
DEFINITION		y79gc12.r1Homo sapiens cDNA clone 154966 5' similar to gb:M80927				
CARTILAGE GLYCOPROTEIN-39 PRECURSOR (HUMAN);.						
ACCESSION		R55530				
NID		9824825				
KEYWORDS		EST.				
SOURCE		human clone-154966 library=Soares breast 2NBHst vector=PT7n3D (Pharmacia) with a modified polylinker host=DH10B (ampicillin resistant) primer=M3Rp1 RSite1-Not I RSite2-Eco RI Adult female. 1st strand cDNA was primed with a Not I - Oligo(dt) primer [5' TGTACCACATCTGAAGTGAGGAGCGGCCCCCTTTTATTATTTTTTTT 3'] double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of modified pRT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 230. Library constructed by Bento Soares and M.Fatima Bonaldo.				
ORGANISM		Homo sapiens				
REFERENCE		Eukaryotes; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonadata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 396) Hillier,L., Clark,N., Dubque,T., Elliston,K., Hawkins,M., Holman,J., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,M., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasius,E., Waterston,R., Williamson,A., Woldmann,P. and Wilson,R.				
TITLE		The Nashu-Merck EST Project				
JOURNAL		Unpublished (1995)				
COMMENT		Contact: Wilson RK Mashu-Merck EST Project Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu High quality sequence stops: 284 Source: IMAGE Consortium, LNL This clone is available royalty-free through LNL; contact the IMAGE Consortium (infoimage.lnl.gov) for further information. Location/Qualifiers 1..396 /organism="Homo sapiens" /clone="154966"				
FEATURES		BASE COUNT	88 a	106 c	109 g	87 t
ORIGIN					6 others	
Query Match		3.9%; Score 65;	DB 134;	Length 396;		
Best Local Similarity		68.3%; Pred. No.1,81e-73;				
Matches		114; Conservative	0; Mismatches	53; Indels	0; Gaps	0;
Db	125	caggtcttgtagtctggtcgttcacagtgtgtctgcatacaaacatggtctact	184			
Oy	49	CAGGTTTCATGGTCTCCTCGTAGATGCATCCATGGGCTCTCTCCAAACTGGTCTGCTACT	108			
Db	185	acacacagctggccacagtaaccgggnangcgatgggaagctgtccacagatgccctgacc	244			
Oy	109	TCACCAACTGGGCCAGATGACAGACAGGGGGAGGCGCTTCTCGCCCAAGACTGGAGCC	168			
Db	245	gntctctgtatccacatcatcatcacgtttgccaatatagaac	291			
Oy	169	CCAGCTTTGACGCCACCTCATCTAGCGCTTGGCTGGCATGACCAAC	215			
RESULT	14	M10705	306 bp	mRNA	EST	26-APR-1996
LOCUS						
DEFINITION		ma7/g07.r1 Soares mouse p3JMF19.5 Mus musculus cDNA 5' similar to				
gb:M80927 CARTILAGE GLYCOPROTEIN-39 PRECURSOR (HUMAN);.						
ACCESSION		M10705				

NID	g1285179	EST.	house mouse.	
KEYWORDS			Mus musculus	
SOURCE			Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae; Murinae; Mus.	
ORGANISM			1 (bases 1 to 306)	
REFERENCE			Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisels,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellander,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.	
AUTHORS			The MashU-HMMI Mouse EST Project	
TITLE			Unpublished (1996)	
JOURNAL				
COMMENT			Contact: Marra M/Mouse EST Project MashU-HMMI Mouse EST Project Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watson.wustl.edu This clone is available royalty-free through LINT; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Seq primer: mob.REGA+ET High quality sequence stop: 172.	
FEATURES				
Source	NCBI g1: 1285179	Location/Qualifiers		
	1..306	/organism="Mus musculus"		
		/note="Vector: pT73D (Pharmacia) with a modified polylinker; Site.1: Not I - Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAGTGGGAGCGGCGCTGATTTTATTTTATTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M. Fatima Bonaldo."		
		/clone_id="Scares mouse p3NMF19.5"		
		/dev_stage="19 weeks"		
		/lab_host="DH10B (ampicillin resistant)"		
BASE COUNT	78 a	76 c	85 g	67 t
ORIGIN				
Query Match	3.6%;	Score 60;	DB 215;	Length 306;
Best Local Similarity	72.1%;	Pred. No. 4.65e-64;		
Matches	98;	Conservative 0;	Mismatches 38;	Indels 0;
			Gaps 0;	
Db	78	agaagcttccttcgctcaacgaagcgaacgagtggtggtatgagacaagagagtg	137	
Qy	985	AGAAGTCCCTACATCTTCGGGACACAGTGGGTGGCTTGATGATGGAGACT	1044	
Db	138	tcaaaacaaggttgggtctcctgaaggaagaaagctgacagagcgcgtgtgtggcac	197	
Qy	1045	TCAAAACCAAGTCACTATCTGAAGCAGAAAGGAGCTGGCGGGCCATGTCGTGGCAC	1104	
Db	198	tggatttgatgatatt	213	
Qy	1105	TGGACTTAGATGACTT	1120	
RESULT	15			
LOCUS	W10705	306 bp	mRNA	EST
DEFINITION	ma77907.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 313884 5'			
ACCESSION	W10705			
NID	g1285179			
KEYWORDS	EST.			

SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Mus.  
REFERENCE 1 (bases 1 to 306)  
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,  
Giesel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,  
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,  
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and  
Waterston,R.  
TITLE The WashU-HMI Mouse EST Project  
JOURNAL Unpublished (1996)  
COMMENT  
Contact: Maria M/Mouse EST Project  
WashU-HMI Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: mob.REGA+ET  
High quality sequence stop: 172.  
Location/Qualifiers  
1..306  
/organism="Mus musculus"  
/note="Vector: p773D (Pharmacia) with a modified  
polylinker. Site.1: Not I; Site.2: Eco RI; 1st strand cDNA  
was primed with a Not I - oligo(dT) primer [5',  
TCTTCCACTCTGAAGTGGAGCGCGCCGATTTTCTTTTCTTTT 3'],  
double-stranded cDNA was size selected, ligated to Eco RI  
adapters (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of a modified p773 vector  
(Pharmacia). Library went through one round of  
normalization to a Cot = 5. Library constructed by Bento  
Soares and M.Falima Bernaldo. RNA was kindly provided by  
Dr. Minoru Ko (Wayne State University)."  
/clone="313884"  
/dev\_stage="19.5 dpc total fetus"  
/lab\_host="DH10B (ampicillin resistant)"  
<1..>306  
MRNA  
BASE COUNT 78 a 76 c 85 g 67 t  
ORIGIN  
Query Match 3.68; Score 60; DB 174; Length 306;  
Best Local Similarity 72.18; Pred. No. 4.65e-64;  
Matches 98; Conservative 0; Mismatches 38; Indels 0; Gaps 0;  
Db 78 agaaggtcccttcgctcgaagagcaccagctgggtggtatgaggaacaaggagatg 137  
||||| ||||| | | ||||| ||||| ||||| ||||| |||||  
QY 985 AGAAGTGCCTCTACATCTCCGACACACCACTGGGTGGCTTGATGATGTGACAGCT 1044  
Db 138 tcaaaacaaggttgggttccttgaagagagaagaagctggcaggagcatcgttgggcac 197  
||||| ||||| | | ||||| ||||| ||||| ||||| |||||  
QY 1045 TCAAAACCAAGGTACGACTACTCAAGACAGAAAGGAGTGGCGGCCGACATGCTGGGCAC 1104  
Db 198 tggatttgatgattt 213  
||||| ||||| ||  
QY 1105 TGGACTTAGTAGTACTT 1120

Search completed: Sun Jun 29 21:08:58 1997  
Job time : 1195 secs.







[illegible]

RESULT	2
ID	R73991 standard; protein; 537 AA.
AC	R73991.
DT	10-JAN-1996 (first entry)
DE	Bovine oviduct specific glycoprotein.
KW	Bovine oviduct specific glycoprotein; recombinant production; BOP.
OS	Bos taurus.
FH	Key
FT	Peptide
FT	/label= sig_peptide
FT	Peptide 19..537
FT	/label= mat_peptide
PN	J07107979-A.
PD	25-APR-1995.
PE	15-AUG-1994; 214227.
PR	19-AUG-1993; JP-227881.
PA	(KINO-) KINOSEI PEPTIDE KENKYUSHO KK.
DR	WP1: 95-190179/25.
DR	N-PSDB; Q90442.
PT	New DNA encoding an oviduct-specific glycoprotein - useful for
PT	recombinant protein production in high quantities.
PS	Claim 4; Pages 7-9; 22pp; Japanese.
CC	Q90442 encodes R73991 bovine oviduct specific glycoprotein (BOP).
CC	The cDNA can be used for the commercial recombinant prodn. of
CC	BOP in high quantities.
CC	Sequence 537 AA;

Query Match:	38.9%	Score 1339;	DB 14;	Length 537;	
Best Local Similarity:	50.1%	Pred. NO. 7.7e-114;			
Matches 190;	Conservative	81;	Mismatches 101;	Indels 7;	Gaps
Db	6	glllvkhhbgaahklycyfnwafszppgasllpdlpdlfctihlyafafasmsnqlyvp	65		
Qy	9	gfmvlmpmgsaaklwcycfnnaqyrogearflpkrldpslcthllyarawfwmhqlst	68		
Db	66	kdpqdeklllyefnfklkernrglktllisigwmfgtlyrftmtstfsnerfswsvtall	125		
Qy	69	twmndet-lyefnfglkkmmnklktllaligwmnstokrfdmataannrotfwnasairfl	127		
Db	126	rthbgfdgldfflypglrspartrwrfvllleelqafkneagldtmrplllisaasgd	185		
Qy	128	krfsfsgldlwmetyfsgosfpaudkerrtllvoolanafoeatstckekrltlaaapag	187		
Db	186	phvvqkayearllgrlldfslvslydldhgswekvtgnspfls-1--pgdps-sa-yam	240		
Qy	188	qtyvadagevdkiaqndfvalmavdphfgsmekxtgtnsnplykrrqeesgaaslindaav	247		
Db	241	nywrtqlyvpekllmqjlytrcthlhllkasgnltaqaavpasbgykxlaqgfiayyeic	300		
Qy	248	qomlqkgpaskllilgmetygrsftllsspsrtfvgaaatpgsgtprtkeggmalyeyvc	307		
Db	301	cfvrrakrkyndgyppafgkygkwygdaisfygaafikrghfganwvrlldldffr	360		
Qy	308	sw-kgatqrqlodokvplrlpfndonwvgddvesskrtvsvytkqkgldgaawmalddlufa	366		

Db 361 gyfcgtgpfplvhtlnll 379  
|:|:|:|:|:|:|  
QY 367 GFSCNQGRYPLIQTLRQEL 385

RESULT	3
ID	R73992 standard; Protein; 718 AA.
AC	R73992;
DT	10-JAN-1996 (first entry)
DE	Murine oviduct specific glycoprotein.
KM	Murine oviduct specific glycoprotein; recombinant production;
KW	MOSP.
OS	Mus musculus.
FX	Key
FT	Peptide
FT	/label= sig_peptide
FT	Peptide 19..718
FT	/label= mat_peptide
PN	U07107979-A.
PD	25-APR-1995.
PF	15-AUG-1994; 2144227.
PR	19-AUG-1993; JP-227681.
PA	(KINO-) KINOMER PEPTIDE KENKUYSHO KK.
DR	WPI; 95-190179/25.
DR	N-PSDB; Q90443.
PT	New DNA encoding an oviduct-specific glycoprotein - useful for
PT	recombinant protein production in high quantities.
PS	Claim 4; Pages 11-14; 22pp; Japanese.
CC	G90443 encodes R73992 murine oviduct specific glycoprotein (MOSP).
CC	The cDNA can be used for the commercial recombinant prodn. of
CC	MOSP in high quantities.
CC	Sequence 718 AA;
CC	

Query March	38.5%	Score 1326; DB 14; Length 718;
Best Local Similarity	49.0%	Pred. No. 1.39e-112;
Matches 130;	Conservative	82; Mismatches 107; Indels 9; Gaps 6;
Db	7	lviImkhsdgtayklyvcyftmwaishrpgpasimphdidpfictllhlfafasmsnqivak 66
QY	11	MYLLMIIPW-GSAAKLVCCYFTTMAQYRQGEARFLPKDLDPSCITLLIYAFAGMTNHOLSTT 69
Db	67	nlgdenlylpefnklkernurelklislgswfgtstftamisltaurekffidsvlsflr 126
QY	70	EWNDET-LYQEFNGLKKKNPKPLKTLTLAGGWNFSTQKFTDMVATANNRQFVNSAIRFLR 128
Db	127	lhgfdgldlffilygllygspphdwtwllflleelqferealltqprillisaavgsip 166
QY	129	KYSPDGLDMEYEGSGSGSERETRTTLYQDANAFQDAQRSGKERLLLSAAPAQ 188
Db	187	sllhtsydaallgrrldifinlysdhlgswsekftghaspfslspedsksaa--y---am 241
QY	189	TYVDGIEVDKIAONLDPVNLMAIDFEGSMKVTGHNHSPILKROESGAASLNDVAHQ 248
Db	242	ywrklygpeadklilngfbygrnflyllkesknglqtasmpaspgkytkeqagflyavcs 301
QY	249	QMLQKGTFASTLLIGMPYGRGSFPLASSSDTRVAPATGSGTPPFLKEGMLAYIVCS 308
Db	302	fvgrakrhwidqyvyvpyafkgkewlgyddtlstfsykamyvkrrehfgganwtldmdvrg 361
QY	309	W-KGATKGRIDQKVPYFRFNDQWVGDDVDSFRTKYSYLKQKLGAGAMWALDDDFAG 367
Db	362	tfcgngpfipvlhll-nellvqtgsnstp 388
QY	368	FSCNQGRRYPLIQTLRQELSLPYLPSGTP 395
RESULT	4	
ID	P81342	standard; protein; 321 AA.
AC	P81342;	
DT	19-OCT-1990	(first entry)
DE	Polypeptide involved in protective mechanisms	
FW	Immune response; cell growth.	
NN	J65032898-A.	



Db 306 gcysvdpstkelisfdtpamistkvswlkqklgsmfw 344





QY 227 PLKROESGAASLNDAVVOQWLKGTPTASKLILMPTGRSFITLASSSDTRVGPAT 286  
 DB 463 hravkqgw--englvdyrqqlasqfmsqewytdgtcteaayvfvpfsgdiltstdasvq 520  
 QY 287 GSGTPEGFTTEGGALAYECS-WKGTQKQ-RIQD-QKVEYIFR-DN-QWVGPDVBSFK 341  
 DB 521 akgyvldkq1991fsweldadn 543  
 QY 342 TKVSYLKQKGLGAMWALDDDD 364

RESULT 14  
 ID R32547 standard; Protein: 238 AA.  
 AC R32547;  
 DT 17-JUN-1993 (first entry)  
 DE HANPY IE-1 gene protein (partial sequence).  
 KW Heliothis armigera nuclear polyhedrosis virus; Ha; NPV;  
 OS Insecticide; immediate early-1.  
 OS Heliothis armigera nuclear polyhedrosis virus.  
 PN MO9303144-A.  
 PD 18-FEB-1993.  
 PF 05-AUG-1992; AU0413.  
 PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
 PI Christian PD, Dail DJ, Gordon KHU, Hanzlik TN, Srisanthia A.  
 DR N-PSDB; Q35982.  
 DR N-PSDB; Q35982.  
 PT Insect virus with reduced capacity to occlude viral particles - used  
 PT for controlling proliferation of insect pests without horizontal  
 PT transmission  
 PS Disclosure: Fig 2: 51pp: English.  
 CC The sequence is that encoded by the Heliothis armigera (Ha)  
 CC nuclear polyhedrosis virus (NPV) immediate early-1 (IE-1) gene  
 CC (partial sequence).  
 CC Sequence 238 AA;  
 SQ

Query Match 7.9%; Score 271; DB 6; Length 238;  
 Best Local Similarity 34.1%; Pred. NO. 1.95e-13;

Matches 58; Conservative 43; Mismatches 51; Indels 18; Gaps 14;

DB 44 erdkfispqwrfevltwkl-fdgididwefpgkqgnupvgdverdnltlyiallgelra 102  
 QY 115 NROTFVNSAI-R-F-L-RKXSFGLDLDWXPSSQS-PAV-DKER----FTLVQDLAN 164  
 DB 103 mldq-vqlgturtleltalsagldkkaav-ncdraqyldkifvmsydfkqawsntdlg 160  
 QY 165 AFQOEAOQTSGERLLLSAAVPAQTYVDAGYEVDKIQNDLFVNLMAYPDHSMEKVT-G 223  
 DB 161 hqataly--gsawkpnepytanvavdalldagrvnp-kklylqvamygrgw 207  
 QY 224 HNSPLTKROESGAASLNDAVVOQWL-QKGTPTASKLILMPTGRSFT 272

RESULT 15  
 ID W02156 standard; Protein: 1046 AA.  
 AC W02156;  
 DT 14-JAN-1997 (first entry)  
 DE Periplasmic chitodextrinase.  
 KW Periplasmic chitodextrinase; periplasmic Beta-N-acetylglicosaminidase;  
 KW Beta-N-acetylglicosaminidase; chitin; oligosaccharide; catabolic;  
 KW catabolism.  
 OS Vibrio furnissii.  
 PN W09625424-A1.  
 PD 22-AUG-1996.  
 PF 13-FEB-1996; U02332.  
 PR 13-FEB-1995; US-386727.  
 PA (Uxjo ) UNIV JOHNS HOPKINS.  
 PI Bassler B, Chlilaru E, Keyhani N, Roseman S, Rowe C;  
 PI Yu C;  
 DR WPI; 96-393335/39.  
 DR N-PSDB; T36387.  
 PT Chitin biosynthetic enzymes end I, exo I and exo II - are

PT periplasmic chito:dextrinase(s), periplasmic beta-glucosidase(s) and  
 PT aryl beta-N-acetyl:gluco:amidase(s), respectively  
 PS Claim 2: Page 68-71, 101pp; English.  
 CC Periplasmic chitodextrinase (W02156), periplasmic  
 CC beta-N-acetylglicosaminidase (W02157) and aryl  
 CC beta-N-acetylglicosaminidase (W02158) can be used to produce chitin  
 CC oligosaccharides with the structure (GlcNAc)<sub>n</sub> where n is 2 or  
 CC higher, by contacting them with soluble chitin. The enzymes are  
 CC encoded by the genes endI, exoI and exoII respectively. They are  
 CC all genes involved in the catabolic pathway of chitin.  
 CC Sequence 1046 AA;  
 SQ

Query Match 6.4%; Score 220; DB 19; Length 1046;  
 Best Local Similarity 31.1%; Pred. NO. 4.80e-09;

Matches 52; Conservative 44; Mismatches 56; Indels 15; Gaps 11;

DB 432 vadqgf-ytmtnadgsinqgletfadsavemmrkyrfqldidrlisidggtppdd 490  
 QY 94 LAIGWNFSQKRTDVAATANNRQTFVNSAIRLRKXSFGLDLD-W--EYPSQSSPA- 149  
 DB 491 tafserraylmshelmrvlrekldvasaqdgqvhymlltaaps-ayllrgmetnavt 549  
 QY 150 VD-KE-RFTLV--ODLANAFQOEAO-TSGKE--RLLSAAVPAQTYVDAGYEVDKIA 201  
 DB 550 qyldvynmsydlhgawndhvgnaalydtgkselagw-nvygtaq 595  
 QY 202 QNDLFVNLMAYPDHSMEKVTGHSPLTKROESGAASLNDAVVO 248

Search completed: Fri Jun 27 16:59:22 1997  
 Job time : 98 secs.







```

QY      68  TTEMNETLVOEFNGLKMPKLTLLAIGWNESTQKFTDMVATANNROTFSVSAIRFL 127
      128  rthbfddglawlyprrr-----dkhfttltkemkeeflkkeag-pqkqjllisaalsag 181
      128  RKTSPDLDDMEYPSGSGSPAYDKERFTTLVODLANAFQEOAQTSGKERLLSAAPAG 187
      182  kvtdesydaklsqhlidflsimtydfhgwrgtqghsp1ffrqeaspdfrfentdyav 241
      188  QTVVDAGYEVDKIAQNNDPFLNMAVDHFGSWEKVTGHNSPLYKROESGAASINVDAAV 247
      242  gymrljlapasaklvmgipctfgsrftlaas-etyvgapisgpgjprftkkaagtlayyeic 300
      248  QOWLOKGTSPASKLTIGMPRTGSRFTLASSSDTRVGAPATGSGTGPPTKEGMLAYEVC 307
      301  dflrgatvthtllggvyackngwvgddgeskxvqylktdqqlagammwldlddfq 360
      308  SW-KGATKORIKODOKVPIRRDNOMVGFDDVESFKTVSTLKQKGLGAGAWMALDDFA 366
      361  gsfccgqlrfpltnakdaa 381
      367  GFSCNQG-RYPLIOTLRQELS 386

RESULT  2
ENTRY   S51327 #type complete
TITLE   heparin-binding glycoprotein 38K - pig
ORGANISM #formal_name Sus scrofa domestica #common_name domestic pig
DATE     07-May-1995 #sequence_revision 07-May-1995 #text_change

ACCESSIONS
REFERENCE S51327
#authors  Shackleton, L.M.; Mann, D.M.; Mills, A.J.T.
#submission submitted to the EMBL Data Library, January 1995
#description Identification of a 38kDa heparin-binding glycoprotein (gp38k) in differentiating vascular smooth muscle cells as a member of a group of proteins associated with tissue remodeling.

#accession S51327
#status     Preliminary
#residues   1-383 ##label SHA
#cross-references EMBL:247803
SUMMARY    #length 383 #molecular-weight 42443 #checksum 2907

Query Match 42.3%; Score 1456; DB 16; Length 383;
Best Local Similarity 50.8%; Pred. No. 1,65e-251;
Matches 197; Conservative 88; Mismatches 94; Indels 9; Gaps 5;

Db 1 mgtrvaqtgfvavljgscaayklvcytswsgyregdgsctfpa1dplcthllysfan 60
QY 1 MVRSVAMAGMVLIMPMGSAALVCYFTWMAQYRGCEARFLPKDLPISCTHLIYAFAG 60
      1 MVRSVAMAGMVLIMPMGSAALVCYFTWMAQYRGCEARFLPKDLPISCTHLIYAFAG 60
      61 tsmneitlwmndctlydntlnknrpnjktllsvggwfgsgfrfkiasntsrftf 120
      61 MTNHQJLSTTMMNDCTLYQEFNGLKMPKLTLLAIGWNESTQKFTDMVATANNROTFSV 120
      121 ksvppftrthbfddglawlyprrr-----dkhfttltkemkeeflkkeag-pqkqjll 174
      121 NSAIRFLKRTSPDLDDMEYPSGSGSPAYDKERFTTLVODLANAFQEOAQTSGKERLL 180
      121 NSAIRFLKRTSPDLDDMEYPSGSGSPAYDKERFTTLVODLANAFQEOAQTSGKERLL 180
      175 sgavsagkvaiddrydlaqisqhlidflsimtydfhgwrgtqghsp1ffrqeaspdfrf 234
      181 SAAVPAGQTVVDAGYEVDKIAQNNDPFLNMAVDHFGSWEKVTGHNSPLYKROESGAAS 240
      235 snadyevsyrljlapasaklvmgipctfgsrftlaas-tdygapaspgjprftkkaagtl 293
      241 LNWDAVAQOWLOKGTSPASKLTIGMPRTGSRFTLASSSDTRVGAPATGSGTGPPTKEGGM 300
      294 layyeicdflgagatvrrpplggvyackngwvgddgeskxvqylktdqqlagammwldld 353
      301 LAIYEVCSW-KGATKORIKODOKVPIRRDNOMVGFDDVESFKTVSTLKQKGLGAGAWMA 359
      354 ldllddfqg-csqpkkefpltnakdaa 381

```

```

QY      360  LDLDDEPAGFSCNQG-RYPLIOTLRQELS 386

RESULT  3
ENTRY   S61550 #type complete
TITLE   BRP39 protein - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE     10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change

ACCESSIONS
REFERENCE S61550
#authors  Morrison, B.W.
#submission submitted to the EMBL Data Library, November 1995
#accession S61550
#status     Preliminary
#residues   1-381 ##label MOR
#cross-references EMBL:X93035
SUMMARY    #length 381 #molecular-weight 43001 #checksum 5107

Query Match 42.2%; Score 1453; DB 14; Length 381;
Best Local Similarity 49.6%; Pred. No. 6,46e-251;
Matches 193; Conservative 89; Mismatches 96; Indels 11; Gaps 9;

Db 1 mgmtraaltgfvavljgscaayklvcytswsgyregvgsflpda1qpfcthllysfan 60
QY 1 MVRSVAMAGMVLIMPMGSAALVCYFTWMAQYRGCEARFLPKDLPISCTHLIYAFAG 60
      1 MVRSVAMAGMVLIMPMGSAALVCYFTWMAQYRGCEARFLPKDLPISCTHLIYAFAG 60
      61 tssdmstlswndesnydklnktrntnltllsvggwkfgkrfseasntertaf 120
      61 MT-NHOLSTTEWNDERLYQEFNGLKMPKLTLLAIGWNESTQKFTDMVATANNROTFSV 119
      121 vrvapflrsygdgdglawlyp-r-lr-dkyfsfllkelaeftevg-pgrkxll 174
      121 VNSAIRFLKRTSPDLDDMEYPSGSGSPAYDKERFTTLVODLANAFQEOAQTSGKERLL 179
      120 VNSAIRFLKRTSPDLDDMEYPSGSGSPAYDKERFTTLVODLANAFQEOAQTSGKERLL 179
      175 lsaalsagkvaiddrydlaqisqhlidflsimtydfhgwrgtqghsp1ffrqeaspdfrf 234
      180 LSAAPVAGQTVVDAGYEVDKIAQNNDPFLNMAVDHFGSWEKVTGHNSPLYKROESGAAS 239
      235 ysnvnyavgymlt1gaaskllmq1pctfgskstlaas-enqljap1sgq1pgrftkkaag 293
      240 SLNWDAVAQOWLOKGTSPASKLTIGMPRTGSRFTLASSSDTRVGAPATGSGTGPPTKEG 299
      294 tlysyelctfklkaevhlsnekvpfatkyngvgyehkesvknxgflkexklagaww 353
      300 MLAYIEVCSW-KGATKORIKODOKVPIRRDNOMVGFDDVESFKTVSTLKQKGLGAGAWW 358
      354 aldllddfqg-tcspkkefpltnakdaa 381
      359 ALDLDDPAGFSCNQG-RYPLIOTLRQELS 386

RESULT  4
ENTRY   I48271 #type complete
TITLE   BRP39 protein - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE     02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change

ACCESSIONS
REFERENCE I48271
#authors  Morrison, B.W.; Leder, P.
#journal   Oncogene (1994) 9:3417-3426
#title     neu and ras initiate murine mammary tumors that share genetic markers generally absent in c-myc and int-2-initiated tumors.

#cross-references MVID:95060797
#accession I48271
#status     Preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues   1-381 ##label RES
#cross-references EMBL:X93035; NID:g1085065; CDS_PID:g1085066

```



[illegible]

KEYWORDS	1-483	#label	KRI
SUMMARY	#residues	1-483	#label KRI
	#cross-references	GB:U01422	
	glycosidase:	hydrolase	
	#length	483	#molecular-weight 52013 #checksum 5938
Query Match	34.3%	Score 1179;	DB 12; Length 483;
Best Local Similarity	42.0%;	Pred. No. 4,64e-197;	
Matches	158; Conservative	101; Mismatches	100; Indels 17; Gaps 11;
Db	19	asphnkvvcyfigawsvyrgngkfdingidptclthiysfvyngkdvkvlpdpwsdipgn	78
Oy	19	GSAAKLVCYFFNMAQYRGGEARFLPKDLDPSCILHLLIYAFAGMTHNOLSTTE--WND--ET	75
Db	79	ldgfgktsllrkhpvsyilmavagqgnagsvpsfgmsdqatreaefgnvvkfllgqyqfd	138
Oy	76	L--YQEFGLKMMPKLTLTALIGGMNFSYQKFDWVAITANNROTFAVNSAIRFLRKYSFD	133
Db	139	gfddwepaqrgspadavkmvklckalkafvq--h---d-yllsaavaapetsak	191
Oy	134	GLDDDMEXPGSGSPAVDWKKEFTTLVQDLANAFQGEAQTSGKERLLLSAAVPAGQTVDA	193
Db	192	sydiaemsqyldfmlntlydfhpgwdhtcmhappssshdsngelklnvkaavkylvlqn	251
Oy	194	GYEVDKIQONIDFYNLMAAYDFHSGMEWYTGHNPSPLYRROESGAASALINDAAVOQMILQ	253
Db	252	gvypkellvgyvpaygkfstlspnsbngklagvsgacgafgytsgnlllynelcemgkag	311
Oy	254	GTPASKLLIGMPTTGRSTTLTSSSDTRGVGAPATISGIFGPFTRKGMALAYEVCSW--KGA	312
Db	312	dwevndehqevpyrvkngvwsfddlaalakakqgikqeglgagmwsietddfkgl-c	370
Oy	313	TKQRID-QK-VPIRFIDNMGWGDVDESFETKYSYLKQKGLGAGMVALDLDFEAFSC	370
Db	371	ge-kypvikaalnsylg	385
Oy	371	NGGRPIILQTLRQELS	386
RESULT	9		
ENTRY	A38221	#type complete	
TITLE	chitinase (EC 3.2.1.14) MFI - nematode (Brugia malayi)		
ORGANISM	#format_name Brugia malayi		
DATE	04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 03-May-1996		
ACCESSIONS	A38221		
REFERENCE	A38221		
#authors	Fuhrman, J.A.; Lane, W.S.; Smith, R.F.; Plessens, W.F.; Perlner, F.B.		
#journal	Proc. Natl. Acad. Sci. U.S.A. (1992) 89:1548-1552		
#title	Transmission-blocking antibodies recognize microfilarial chitinase in brugian lymphatic filariasis.		
#cross-references	MUID:92179220		
#accession	A38221		
#status	preliminary; not compared with conceptual translation		
#molecule_type	nucleic acid; protein		
#residues	1-504 #label FWH		
#cross-references	NCBI:85345		
#note	sequence extracted from NCBI backbone		
KEYWORDS	glycosidase; hydrolase		
SUMMARY	#length 504 #molecule-weight 55971 #checksum 7343		
Query Match	31.4%;	Score 1080;	DB 12; Length 504;
Best Local Similarity	42.5%;	Pred. No. 1.02e-177;	
Matches	157; Conservative	95; Mismatches	102; Indels 15; Gaps 11;
Db	27	cyttwagrydggkflipgnlpgllchliylafakvdelqskpfewndetewskgmys	86
Oy	26	CYTFWMAQYRGGEARFLPKDLDPSCILHLLIYAFAGMTHNOLSTT--EWNDE-T-----LYQ	78
Db	87	avtkirenpqklkllsyygynfsafstlgaakaqkterfiksaafltknmfdgfdl	146
Oy	79	EPNGILKKNPKLTKTLTALIGGMNFSYQKFDWVAITANNROTFAVNSAIRFLRKYSFDGDL	138

Db	147	weyr-v-gv-a-e-eh-aklyeamktafveaktsqgrlll1taavaagxgldgsyve	200
Oy	139	WEYPSQSPVADKRRFTTLVQDLANAEQDAQTSIGKERLLLSAIVAGQRYVAGXEVD	198
Db	201	sigkfdllfmsydlbhswekndvlbklipktgsvsglftfteadqwasgmpke	260
Oy	199	KIAQNLDFVNLMAADFHSMEKVTGHNSPDLTKROEESGAASLNVDAVQOMLOKQYFAS	258
Db	261	klllg1pmayagwclndpnsctatgaasrpsaaktimpagcaasyelckylkegkety	320
Oy	259	KLLIGMPYRGSFTLIASSSDTRVGNAPARGSOTPGPFTRKEGMLAYEVCSS-KKATQRI	317
Db	321	hgeyvgaymvgkdgwyrygneetlrkmlkylekyggaftwaldfdctgkscgkyp	380
Oy	318	ODOKV-PYIFPDNOMVGDDVESEFTKYSYLKOKGLGAMWALDLDPDFAGFSCNGRYP	376
Db	381	llna1sel 389	
Oy	377	LIQTLROEL 385	
RESULT	10		
ENTRY	A56596	#type complete	
TITLE	chitinase (EC 3.2.1.14) - tobacco hornworm		
ORGANISM	formal_name Manduca sexta #common_name tobacco hornworm		
DATE	11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 03-May-1996		
ACCESSIONS	A56596		
REFERENCES	A56596		
Authors	Kramer, K.J.; Corpuz, L.; Choi, H.K.; Muthukrishnan, S.		
#journal	Insect Biochem. Mol. Biol. (1993) 23:691-701		
#title	Sequence of a cDNA and expression of the gene encoding epidermal and gut chitinases of Manduca sexta.		
#cross-references	MUID:93357793		
#accession	A56596		
#status	preliminary		
#molecule_type	mRNA		
#residues	1-554 #label KPA		
#cross-references	NCBI:136417, NCBI:136418		
#experimental_source	larvae		
#note	sequence extracted from NCBI backbone glycosidase: hydrolase		
KEYWORDS	#length 554 #molecular_weight 62203 #checksum 4659		
SUMMARY			
Query Match	29.8%; Score 1025; DB 12; Length 554;		
Best Local Similarity	38.5%; Pred. No. 5,30e-167;		
Matches 161; Conservative	106; Mismatches 126; Indels 25; Gaps 19;		
Db	17	vgdsrarivcvgfsmwayryprgygjedipvekcchlysflygtegneasevllldpel	76
Oy	16	IPWGSAAFLVCFITWMAOYRGSEARFLPKDLPCLCHTLIYAFAGMNN-H-Q-LST-TW	71
Db	77	dvdkngfnfcsfshsbvskfmavvgwaegsskyslmvaqsktsmfsfrsvsflky	136
Oy	72	N-DETLXOEFNGLKKMNPKLTLLAIGGMNFSSTOKFTDMVATANNROTFFVNSAIRFRKY	130
Db	137	dtdgdldlweyrgaadsrgsfcdkflylqglrnf---lrv-gygel-teayplan	191
Oy	131	SFDGIDLWETPGS--QSSPAVDKERFTTLVQDLANAEQDAQTSIGKERLLLSAIVAGQ	188
Db	192	flrimgyhnvelcge1de1hmsydlrgnwagfadvnarlkykrphdwayeklnvdq1h	251
Oy	189	TYVDAGYEVDKIAQNLDFVNLMAADFHSMEKVTGHNSPDLTKROEESGAASLNVDAVQ	248
Db	252	lwekgcgsnklvnglpyrgstfcsagmnnyglftlnkaagsgdpyatnatqfayy	311
Oy	249	QWLQGTGTPASKRLILCMPTYGRSFTLIASSSDT-RVGA--P-ATGSGTRGPFTRKEGMLAY	304
Db	312	elctevvdsgsgkfwegkcpayaygltwgyedprsvelkmvylkcdqyigantta	371
Oy	305	EVCs--WKG-A--TKORLQDQVPIIFPDNOMVGDDVESEFTKYSYLKOKGLGAGAWMA	359
Db	372	ldmdfdggl-cge-knplikllhkms--syvpphtentpfbewarpspspsdpe	426

QY	360	LDLDFAFSCNQGKRYPIQTLROELSLUPY-LPSPGTRPELVKPGQSEPEHGPSPOQ	416						
RESULT	11								
ENTRY	A38368	#type complete							
TITLE	chitinase (EC 3.2.1.14) precursor	- Bacillus circulans							
ORGANISM	#formal_name	Bacillus circulans							
DATE	28-Jun-1991	#sequence_revision	28-Jun-1991 #text_change						
ACCESSIONS	A38368								
REFERENCE	A38368								
#authors	Watanabe, T.; Suzuki, K.; Oyanagi, W.; Ohnishi, K.; Tanaka, H.								
#journal	J. Biol. Chem. (1990) 265:15659-15665								
#title	Gene cloning of chitinase A1 from Bacillus circulans WL-12 revealed its evolutionary relationship to Serratia chitinase and to the type III homology units of fibronectin.								
#cross-references	MUID:9036876								
#accession	A38368								
##status	preliminary								
##molecule-type	DNA								
##residues	1-699	#label	WAT						
##cross-references	GB:J05599								
CLASSIFICATION	#superfamily	fibronectin type III repeat homology							
KEYWORDS	glycosidase; hydrolase								
SUMMARY	#length	699	#molecular-weight	73677	#checksum	6757			
Query Match	20.3%	Score	700;	DB	10;	Length	699;		
Best Local Similarity	40.4%	Pred. No.	2,24e-104;						
Matches	120;	Conservative	72;	Mismatches	85;	Indels	20;	Gaps	18;
Db	144	qlnkrlkqnpnlkllslsvgwtswn-rfsdtaatacrevfansavdflkynfdgvdld	202						
QY	79	EFNGIKKNPKIKTLAIGGNFSTOFITDWAFAANNKQFVNSAIRFLKRYSTDGLDLD	138						
Db	203	weypvgglldgnskrpdxnyllllakireklidaagvgdk-kyllltlsgasatya-a	260						
QY	139	WEYP-GS-QG-S-PAYDKERFTLVODLANAFQOEKQTSKEKRLLSAAVPMQITVDA	193						
Db	261	ntelakiaivdwlmtydfingawqklsahnaplnjdpaasaagvdpdantfnvaagqg	320						
QY	194	GYEVXDKNQNDPYNLMAVDHSGMENVYTGNSPL-YKRQE-ESGAA-A-SLVNDAVQO	249						
Db	321	hlldagvpaakivlvvpyfgrwd-gcaqaqgn-qgvyqtcggssvgtweaasfdydlan	378						
QY	250	WLQKTPPAKSKILLMPYGRSFTLASSSDRYVGAPATGSGTPGCFTRKEGMLAYEY-CS	308						
Db	379	ylnkgyrrywndcakvpylynaankffisyddesvgyktaylxksyglgmfwel	435						
QY	309	W-K-GATKQRIODOKVPIFRD-NQW-VGPDVDESFRKTVSKYKQKGLGAMVWAL	360						
RESULT	12								
ENTRY	S51591	#type complete							
TITLE	chitinase (EC 3.2.1.14)/lysozyme (EC 3.2.1.17) P2 precursor								
ORGANISM	#formal_name	Nicotiana tabacum	#common_name	common tobacco					
DATE	15-Jul-1995	#sequence_revision	08-Sep-1995	#text_change					
ACCESSIONS	S51591;	S51633;	S43119						
REFERENCE	S51591								
#authors	Helicz, T.; Segond, S.; Kaufmann, S.; Geoffroy, P.; Prasad, V.; Brunner, F.; Frilley, B.; Legrand, M.								
#journal	Mol. Gen. Genet. (1994) 245:246-254								
#title	Molecular characterization of a novel tobacco pathogenesis-related (PR) protein: a new plant chitinase/lysozyme.								
#accession	S51591								
##molecule-type	mRNA								
##residues	1-378	#label	HEI						
##cross-references	EMBL:X78325								

```

##experimental_source cv. Samsun NN
#accession S51632
#molecule_type protein
##residues 31-36:87-112:252-275:282-305:337-371 ##label HW
CLASSIFICATION #superfamily Streptomyces chitinase chl40
KEYWORDS glycosidase; hydrolase
FEATURES
1-25 #domain signal sequence #status predicted #label SIG\
26-371 #product chitinase/lysozyme P2 #status predicted #label
371-378 #domain carboxyl-terminal propeptide #status
#experimental #label PRO
SUMMARY #length 378 #molecular-weight 42019 #checksum 6689

Query Match 17.3%; Score 594; DB 5; Length 378;
Best Local Similarity 32.7%; Pred. No. 2,53e-84;
Matches 106; Conservative 87; Mismatches 108; Indels 23; Gaps 18;

Db 43 mldstlflhlfcafdlmpsqnqllispengds-fsgftstvrknpkvkflsiagr 101
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 44 KDLDPSELCTHLIAFAAGMT--NHQLSTEMNDETLYQEFNG-LKKNNPKLTLLAI-GGM 99
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 102 a-dttayglnaqpnrskfidsitlarqfghjdlwep--l-s-atdmtnlqll 156
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 100 NFSTQKFTDMVATANNKQTFVNSAIRFLRKYSFDGLDMEXPGSGSPAVDKERTTLV 159
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 157 newrtaimneansgrnaallltaavysprvnglnyvesvarnlwlmaydfyqpw 216
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 160 QDLANAFQOEAGTSGKERLLLSAAYVAGQTYDAGTEVDKIDNLPVNLMAVDFFGS-W 218
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 217 spsqtnhaqlfdpvnh--lsgs--dg-inawigavptkklvlpfygawrlvnpn 270
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 219 EKV-TEHNSPLYKROESGAASLANVDAVQMLKQKTPASKILLMPYGRSFTLASS 277
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 271 indlrpaagksnvga-vddgsm-tyritrdyivgsrattvnaivdycysgmwisy 328
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 278 DTRVGAAPAGSGTPGPFTEGGLAYEVCMSWGATKOR-IDOK-VP-IFRDNOMVGF 334
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 329 ddtgsyrvnkvnvkgfqllygfaw 352
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 335 DVESFRTKVSYLKQKGLGAMVW 358
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 13
ENTRY S47133 #type complete
TITLE chitinase (EC 3.2.1.14) - fungus (Trichoderma harzianum)
ORGANISM #formal_name Trichoderma harzianum
DATE 23-Nov-1994 #sequence_revision 03-Aug-1995 #text_change
10-May-1996
ACCESSIONS S47133
REFERENCE S47133
#authors Carlsio, C.; Gutierrez, A.; Jimenez, B.; van Montagu, M.;
#submission Herrera-Estrella, A.
#description submitted to the EMBL Data Library, June 1994
#endochitinase Characterization of ech-42, a Trichoderma harzianum
#accession S47133
#status preliminary
#molecule_type DNA
##residues 1-424 ##label CAR
##cross-references EMBL:X79381
GENETICS
#introns 48/2: 81/2: 98/1
CLASSIFICATION #superfamily Streptomyces chitinase chl40
KEYWORDS glycosidase; hydrolase
SUMMARY #length 424 #molecular-weight 46298 #checksum 8773

Query Match 17.0%; Score 585; DB 5; Length 424;
Best Local Similarity 36.4%; Pred. No. 1.24e-82;
Matches 103; Conservative 68; Mismatches 95; Indels 17; Gaps 13;

Db 112 glfk-lkkanrnlykmslsgwtwt-nfpaastdanknfkaltlmkdgfgldy 169
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

```

Qy 78 QEFNLKKMNPXKLTLLAIGGNFSTQKFTDMVATANNKQTFVNSAIRFLRKYSFDGL 137
Db 170 dweyp-addtqttnwvlllkelrsgldayaay-apy-yhflslsaapeghy-sflm 225
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 138 DWYFSGSPAVDKERTTLVQDLANAFQOEAGTSGKERLLLSAAYVAGQTYDAGYEV 197
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 226 sdlgavlgyvnlmaydyagswssyghdanlfanpnssp-ytltdqakdyikgyra 284
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 198 DLIAQNLDPVNLMAVDFFGSMEKVTGHNSPLYKROESGAASLANVDAVQMLKQKTPA 257
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 285 skivlmpiygraf--estga--lqcyyslgs-gsw--englwdy-kvlpkagatgyd 336
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 258 SKLIIGMPYGRSFTLASSSDTRVGAAPATGSGTGPFTKEGGLAYEVCMSWGATKO-- 315
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 337 swagayysvdpsskelisfdpdmintkvsylknlgsgmfw 379
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 316 RIQDOKVPIFRDNOMVGFDDVESFRTKVSYLKQKGLGAMVW 358
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 14
ENTRY J01975 #type complete
TITLE chitinase (EC 3.2.1.14) 1 precursor - imperfect fungus
ORGANISM #formal_name Aphanocladium album
DATE 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change
03-May-1996
ACCESSIONS J01975
REFERENCE J01975
#authors Blaiseau, P.L.; Tafay, J.F.
#journal Gene (1992) 120:243-248
#title Primary structure of a chitinase-encoding gene (chl) from
the filamentous fungus Aphanocladium album: similarity to
bacterial chitlinases.
#accession J01975
#molecule_type mRNA
##residues 1-423 ##label BLA
##cross-references GB:X64104
COMMENT This enzyme is essential for the degradation of insoluble chitin.
GENETICS
#gene chl1
#introns 47/2: 80/2: 97/1
CLASSIFICATION #superfamily Streptomyces chitinase chl40
KEYWORDS glycosidase; hydrolase
FEATURE 1-34
SUMMARY #length 423 #molecular-weight 46072 #checksum 8802

Query Match 16.8%; Score 579; DB 5; Length 423;
Best Local Similarity 36.6%; Pred. No. 1.66e-81;
Matches 102; Conservative 64; Mismatches 95; Indels 18; Gaps 16;

Db 115 lkkqrmnkvmjslsgwtwt-nfpaassaatrttfgsavgfinkdvgfdgidwep 173
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 83 LKKMNPXKLTLLAIGGNFSTQKFTDMVATANNKQTFVNSAIRFLRKYSFDGLDMEXY 142
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 174 -datagqumvlllgavseidsyaay-akg-hhflslsaapepdylnklfa-e-lg 228
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 143 GSGSPAVDKERTTLVQDLANAFQOEAGTSGKERLLLSAAYVAGQTYDAGYEVDDA 201
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 229 kvldylnmaydyagswssygtghdanlfanpnpnatp-yntddavqaylmgvpankv 287
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 202 QNLDPVNLMAVDFFGSMEKVTGHNSPLYKROESGAASLANVDAVQMLKQKTPASKLI 261
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 288 lmpiygrsfq-qt--eg-igkpynglgs-gsw--englwdykalpk-agatvkcdtdak 339
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 262 LGMPYGRSFTLASSSDTRVGAAPATGSGTGPFTKEGGLAYEVCMSWGATKORIDOK 321
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 340 gcyysvdpsskelisfdpdmintkvsylkylgsgmfw 378
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 322 VPIFRD-NOMVGFDDVESFRTKVSYLKQKGLGAMVW 358
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

RESULT	15
ENTRY	

ENTRY TITLE	#type complete
S51369 chitinase - fungus ( <i>Trichoderma harzianum</i> )	

ORGANISM	#formal_name	Trichoderma harzianum
DATE	15-Jul-1995	#sequence_revision 19-Apr-1996
		#text_change

03-May-1996  
CE1360

ACCESSIONS	S51369
REFERENCE	S51369

#authors García, I.; Lora, J.M.; de la Cruz, J.; Benítez, T.; Llobell, A.; Píntor-Toro, J.A.  
#journal Curr Biol 2004, 14, 27-33

#journal	Curr. Genet.	(1994)	27:83-89
#title	Cloning and characterization of a chitinase (CHIT42) cDNA		

from the mycoparasitic fungus *Trichoderma harzianum*.  
#accession S51369

```
##status      preliminary
##molecule_type DNA
##molecule_type DNA
```

```
##residues      1-423  ##label GAR
#length 423  #molecular-weight 46056  #checksum 2000
```

Query Match 16.4% Score 564: DB 12: Length 423:

Best Local Similarity 36.0%; Pred. No. 1.07e-78;  
Matches 102; Conservative 70; Mismatches 94; Indels 17; Gaps 13

Db 111 qlfk-vkangrlkvlslsgwtwsc-nfpsaasrdanrkntfakta1tfmxcwgfddq1d1 168

```
09      | : : || | | : : |   : | | : : :||| :  
78 QEENGKMKMPKLKTLLAIGGWNFSTQFTDMVAFNRRQTFFVNSAIRFLRKXSPDGLDL 137
```

Db 169 dweyp-adatqasm1111kvevrsqrdayaay-aps-yhfl1t1aapagkany-sk1r1 224

```
Oy      |||| :...| :: : | : : : ||| :  
        DMEYPGSGSPAVDKERETTLVQDLANFQQEAGTSGKERLLLSA VPAGQTYYDAGEY 197
```

Db 225 adlqgvldylnlmaydyagsfspitghdanlfmnpnpatp-fntdsavkdylnngvpa 283

Oy 198 DKIAQNLDENVNLMAAYDFGHSWEKVTGHSNPLYKRRQESGAALSLNVDAAVQCMQKGTPTA 257

Db 284 nkivlgmp4ygrsf--qntag--lgqfyngvgs-gsw--eaqlwdqkalpk-agatvqyd 335

0y 258 SKILGMPYGRSFTLASSDTRVGAPATGSGTPEGFTKEGMLAYEEVCSNKGATKQRI 317

336 svakgyysynsatkelisfcdpdmnrkwaylksjlgsgsmfw 378

0Y 318 QDQKVPYIFRD--NQWVGFDVSEFKTKVSYLKQKSGAMVW 358

2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841

Search completed: Fri Jun 27 16:57:26 1997  
Job time : 72 secs.

Search completed: Fri Jun 27 16:57:26 1997  
Job time : 72 secs.

100







[illegible]

RESULT	2	STANDARD;	PRT;	504 AA.
ID	CHIT1_BROMA			
AC	P29030;			
DT	01-DEC-1992 (REL. 24, CREATED)			
DT	01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)			
DT	01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)			
DE	ENDOCHITINASE PRECURSOR (EC 3.2.1.14) (MFI ANTIGEN).			
OS	BORGIA MALAYI.			
OC	ARABIDOPSIS, METAZOA; ACCELLOMATES; NEMATODA; SECERNENTEAE; SPIRURIDA.			
RN	[1]			
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RX	MEDLINE; 92179220.			
RA	FUHRMAN J.A., LANE W.S., SMITH R.F., PLESSSENS W.F., PERLER F.B.;			
RL	PROC. NATL. ACAD. SCI. U.S.A. 89:11548-11552(1992).			
CC	-1- FUNCTION: THE MFI ANTIGEN IS A MICROFILARIAL CHITINASE, WHICH MAY			
CC	FLUARA OR IN ITS MOSQUITO VECTOR DURING PARASITE DEVELOPMENT AND			
CC	TRANSMISSION.			
CC	-1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE BETA-1,4 LINKAGES OF			
CC	N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.			
CC	-1- PTM: O-GLYCOSYLATED.			
CC	-1- KNOWN TO BIND CALCIUM.			
CC	-1- DEVELOPMENTAL STAGE: THE APPEARANCE OF THE MFI ANTIGEN CORRESPOND			
CC	WITH THE ONSET OF THE PARASITE'S ABILITY TO INFECT THE MOSQUITO.			
CC	-1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GYCOSYL			
CC	HYDROLASES).			
DR	EMBL; M73689; G156064; -.			
DR	PIR; A38221; A38221.			
DR	PROSITE; PS01095; CHITINASE_18.			
KW	HYDROLASE; GLYCOSIDASE; CHITIN DEGRADATION; SIGNAL; ANTIGEN;			
KW	REPEAT; GLYCOPROTEIN; CALCIUM-BINDING.			
FT	SIGNAL	1	22	
FT	CHAIN	23	504	ENDOCHITINASE.
FT	DOMAIN	23	400	CATALYTIC.
FT	DOMAIN	401	450	SER/THR-RICH (LINKER).
FT	DOMAIN	407	448	3 X 14 AA APPROXIMATE TANDEM REPEATS.
FT	ACT_SITE	148	148	PROTON DONOR (BY SIMILARITY).
SO	SEQUENCE	504 AA;	55571 MW;	4DA/5EB CRC32;

Query Match	31.48;	Score 1080;	DB 2;	Length 504;
Best Local Similarity	42.58;	Pred. No. 4.93e-225;		
Matches 157;	Conservative	95;	Mismatches 102;	Indels 15;
				Gaps 11;

db 27 cyttwagyrldgegrkflipnlpnglcthllyafakvdeygdskpfendedeitwskgmys 86  
||:|||||:::||:||||:| ||||| :|  
QY 26 cyftnwagyrqgearfLpkdlDpslcThLyfAGMTNQLSTT-ENDE-T-----LYQ 78

```

Db      87 avtklretnpqlkvllysgygnfgsalfitgiaksagctefzifsalafrkmfiofdld 146
Qy      79 EFNGILKKNNPKLKLTLAIGGNFSTQKFTMTVAANNRQFVSSALFRFLKRYSFSDGUDJ 138
Db      147 weyp-v-gv-a-e-eh-aklveamktafeeaktsqqrllltaavsaqgttdgsyne 200
Qy      139 WEYEGSOGSPYVNDERRTTLVQDLAANFOQDAOTSGERILLISAAVPAQGTYYDAGTEYD 198
Db      201 slgkfdllfmsydlhngsewknvdlhgklhpketgsyigifntefaadyasakmpxe 260
Qy      199 KIAQNLDFVNLMAVDFHSGMEKVTGHNSPLYKQGEESGAASLNVDAAYQWLOKGRPNs 258
Db      261 kiliigipmyagawcltdnpsetajaaasirpssasklnpagtasyweickylkeggketv 320
Qy      259 KLIIEMPYGKSFYLASSSDTRVAGAPATGSGTGPFFKKEGGMLAYVEVCSM-KGATKQRI 317
Db      321 hgeyggaymmvkgdqvwydydneetlrlkmkvlkekyggafitwaldfdidfyskscgkpypp 380
Qy      318 QDQKRY-PIYFDRDNQWVGFDDVESEKTVSYLKKOKGLGAGAMVALDDDDFEGSCNOGRYP 376
Db      381 llnaisael 389
Qy      377 LIQTLROEL 385

```

RESULT	3	STANDARD;	PRT;	554 AA.
ID	CHIT MANSE			
AC	P36362;			
DT	01-JUN-1994 (REL. 29, CREATED)			
DT	01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)			
DT	01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)			
DE	ENDOCHITINASE PRECURSOR (EC 3.2.1.14).			
OS	MANDUCA SEXTA (TOBACCO HAWKMOTH) (TOBACCO HORNWORM).			
OC	EMKAROTA; METAOTA; ARTHROPODA; INSECTA; LEPIDOPTERA.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE; 93357793.			
RA	KRAMER K. J., CORPUZ L., CHOI H. K., MUTHUKRISHNAN S. ;			
RL	INSECT BIOCHEM. MOL. BIOL. 23:651-701(1993).			
CC	-1- FUNCTION: DIGEST CHITIN IN THE EXOSKELETON DURING THE MOLTING PROCESS.			
CC	-1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE BETA-1,4 LINKAGES OF			
CC	N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.			
CC	-1- DEVELOPMENTAL STAGE: HIGH LEVELS SEEN IN THE EPIDERMIS ON DAY 0,			
CC	BUT RAPIDLY DISAPPEARS AND IS UNDETECTED ON DAYS 1-4 OF FIFTH			
CC	INSTAR. IT REAPPEARS ON DAY 5 AND PEAKS ON DAY 7 AFTER WHICH A			
CC	RAPID DECLINE IS SEEN. IN THE GUT IT IS DETECTED ON DAY 6 WITH LOWER			
CC	LEVELS SEEN ON DAYS 0, 7 AND 8.			
CC	-1- TISSUE SPECIFICITY: EPIDERMIS AND GUT.			
CC	-1- SUBCELLULAR LOCATION: SECRETED.			
CC	-1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL			
CC	HYDROLASES).			
DR	EMBL; U02270; G406049; -.			
DR	EMBL; S64757; -; NOT_ANNOTATED.CDS.			
DR	HSSP; P07254; ICTN.			
DR	PROSITE; PS01095; CHITINASE.18.			
KW	HYDROLASE; GLYCOSIDASE; CHITIN DEGRADATION; SIGNAL; GLYCOPROTEIN.			
FT	SIGNAL	1	19	POTENTIAL.
FT	CHAIN	20	554	ENDOCHITINASE.
FT	DOMAIN	396	453	SER/THR-RICH.
FT	ACT_SITE	146	146	PROTON DONOR (BY SIMILARITY).
FT	CARBOHYD	85	85	POTENTIAL.
FT	CARBOHYD	303	303	POTENTIAL.
FT	CARBOHYD	407	407	POTENTIAL.
FT	CARBOHYD	545	545	POTENTIAL.
SO	SEQUENCE	554 AA;	62203 MW;	FA87/F8AD CRC32;

Query Match	29.8%;	Score 1025;	DB 2;	Length 554;
Best Local Similarity	38.5%;	Pred. No. 1,84e-211;		
Matches	161;	Conservative	106;	Mismatches 126;
			Indels	25;
			Gaps	19;

```
Db      17 vqdsrrivcyfsnwavyrrpgvrgryjedipvekcthllysflygtegnsevalldpel 76
       : |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
       : |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

QY 16 IPMSAAKVCYFTNNMOYRGEARFLPKDLPSCITHLIYAFAGMTN-H-Q-LST-TW 71  
 DB 77 dvdkngfnfslspsvkfwagvaggsskyswmaqktsmfsfrsvsflkky 136  
 QY 72 N-DETLVQEFNGKLKMPKLTLLAIGMNFSTOKFTDMVATANNROTFVNSAIRFLRKX 130  
 DB 137 dffgldldkweygsaadggsgfsdkflylvqelrraf---lrv-gygwel-taavplan 191  
 QY 131 SFDDLDDWMEYPGS--OGSPAVDKEKERTTLVQDLANAFQOEAOQTSCKERLLLSAAVPAGQ 188  
 DB 192 frlmeqyhyvelcgealdahmsydlrignwagfadvhsplykrphddqwayeklvndgln 251  
 QY 189 TYVDAGLEVQKIQNDLFVNLMAYDFHSGMEKVTGHSNPLKROBESGAASLAVDAVQ 248  
 DB 252 lweekgcpnklvvglyfygrsfclssagnnnylgfinkwaagggdparytnatgfwayy 311  
 QY 249 QWLQKGPASKLILGMPTYGRSFTLASSSDR-RVGA--P-ATGSGTGPFTKEGMLAYY 304  
 DB 312 elctevdkddsgwtkkvddegkcpaykgytgywgyedprsvetlmmvlykqgylgantwa 371  
 QY 305 EVCS--WKG-A--TKQRIQDQKVPYIFRDNQWGFDDVESKRTVSYLKQGLGAMVWA 359  
 DB 372 ldmddfggl-cge-knpdklkhkms-sylvpphtcentpdpwarpstpsdpse 426  
 QY 360 LDLDDFGFCNGGRYPLIOTLRQELSLPY-LPSCITPELVKPKGQPSRPHGSPQ 416  
 RESULT 4  
 ID CH11\_BACCI STANDARD; PRT; 699 AA.  
 AC P20533;  
 DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1991 (REL. 31, LAST SEQUENCE UPDATE)  
 DE CHITINASE A1 PRECURSOR (EC 3.2.1.14).  
 GN CHIA1.  
 OS BACILLUS CIRCULANS.  
 OC PROKARYOTA: FIRMITICES: ENDOSPORE-FORMING RODS AND COCCI: BACILLACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-WL-12;  
 RX MEDLINE: 90368776.  
 RA WATANABE T., SUZUKI K., OYANAGI W., OHNISHI K., TANAKA H.;  
 RL J. BIOL. CHEM. 265:15659-15665(1990).  
 RN [2]  
 RP MOTAGENESIS.  
 RC STRAIN-WL-12;  
 RX MEDLINE: 93366760.  
 RA WATANABE T., KOHORI K., MIYASHITA K., FUJII T., SAKAI H.,  
 RL UCHIDA M., TANAKA H.;  
 CC J. BIOL. CHEM. 268:18567-18572(1993).  
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE BETA-1,4 LINKAGES OF  
 N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.  
 CC -1- SIMILARITY: CONTAINS 2 FIBROECTIN TYPE III-LIKE DOMAINS.  
 CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
 HYDROLASES).  
 DR EMBL: M57601; G442688; -.  
 DR PIR: A38368; A38368.  
 DR HSSP: P07254; 1CTN.  
 DR PROSITE: PS01095; CHITINASE.18.  
 KW HYDROLASE; GLYCOSIDASE; CHITIN DEGRADATION; SIGNAL; REPEAT.  
 FT SIGNAL 1 41  
 FT CHAIN 1 41  
 FT DOMAIN 42 699 CHITINASE A1.  
 FT DOMAIN 42 460 CATALYTIC.  
 FT DOMAIN 549 549 FIBROECTIN TYPE-III (R-1).  
 FT DOMAIN 560 644 FIBROECTIN TYPE-III (R-2).  
 FT ACT\_SITE 204 204 PROTON DONOR (PROBABLE).  
 FT MUTAGEN 200 200 D->N: DECREASE IN ACTIVITY.  
 FT MUTAGEN 200 200 D->E: NO CHANGE IN ACTIVITY.  
 FT MUTAGEN 204 204 E->D,Q: LOSS OF ACTIVITY.  
 SQ SEQUENCE 699 AA; 73677 MW; 5A1777CC CRC32;  
 Query Match 20.3%; Score 700; DB 2; Length 699;  
 Best Local Similarity 40.4%; Pred. No. 3.64e-132;

Matches 120; Conservative 72; Mismatches 85; Indels 20; Gaps 18;  
 DB 144 qlnkklctnplktlsvvgwtvan-rfsdvaataatrewfanaavglkkyndgvdld 202  
 QY 79 EFNGLKMPKLTLLAIGMNFSTOKFTDMVATANNROTFVNSAIRFLRKXSPDGLD 138  
 DB 203 wepvsaglgnskrpdknqylllekirekidaaavqk-kylltllsagaaty-a 260  
 QY 139 WEYR--GS-YG-S-PAYDKERTTLVQDLANAFQOEAOQTSCKERLLLSAAVPAGQ 193  
 DB 261 ntelaklaavdwlnmtcylngagwksahnaplndpaaasagvdpantfnvaagq 320  
 QY 194 GYEDKIAQNDLFVNLMAYDFHSGMEKVTGHSNPL-TRQD-ESGAA-A-SLNDAAVQ 249  
 DB 321 hldagvpaaklvlyvpygrwd-gcaqagn-gqyqctgssvgtwaeasfdfydlean 378  
 QY 250 WLQKGPASKLILGMPTYGRSFTLASSSDR-RVGA-PATGSGTGPFTKEGMLAYY-ES 308  
 DB 379 ylnknglyrmdtakpylynaankfisyddaesvglytalyksqglgamtwa 435  
 QY 309 W-K-GATKORIDQKVPYIFRD-NQW-VGFDVESKRTVSYLKQGLGAMVWA 360  
 RESULT 5  
 ID CH11\_APHAL STANDARD; PRT; 423 AA.  
 AC P32470;  
 DT 01-OCT-1993 (REL. 27, CREATED)  
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
 DE CHITINASE 1 PRECURSOR (EC 3.2.1.14).  
 GN CHIA1.  
 OS APHANOCCLADIUM ALBUM.  
 OC EUKARYOTA: FUNGI: DEUTEROMYCOTINA (IMPERFECT FUNGI).  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ETHM 483;  
 RX MEDLINE: 93013040.  
 RA BLAISEAU P.-L., LAPAY J.-F.;  
 RL GENE 120:243-248(1992).  
 RN [2]  
 RP SEQUENCE OF 35-57.  
 RX MEDLINE: 92136437.  
 RA BLAISEAU P.-L., KUNZ C., GRISON R., BERTHEAU Y., BRIGOO Y.;  
 RL CURR. GENET. 21:61-66(1992).  
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE BETA-1,4 LINKAGES OF  
 N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.  
 CC -1- SUBCELLULAR LOCATION: SECRETED.  
 CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
 HYDROLASES).  
 DR EMBL: X64104; G429026; -.  
 DR PIR: J01975; J01975.  
 DR HSSP: P07254; 1CTN.  
 DR PROSITE: PS01095; CHITINASE.18.  
 KW HYDROLASE; GLYCOSIDASE; CHITIN DEGRADATION; SIGNAL; ZMOGEN.  
 FT SIGNAL 1 22  
 FT PROPEP 23 34 POTENTIAL.  
 FT ACT\_SITE 35 423 CHITINASE 1.  
 FT ACT\_SITE 171 171 PROTON DONOR (BY SIMILARITY).  
 SQ SEQUENCE 423 AA; 46058 MW; 884565E6 CRC32;  
 Query Match 16.8%; Score 579; DB 2; Length 423;  
 Best Local Similarity 36.6%; Pred. No. 3.08e-103;  
 Matches 102; Conservative 64; Mismatches 95; Indels 18; Gaps 16;  
 DB 115 lkknrmkxlmislygtwat-nfpaasaatrkktagavgmkmkgfdgldldwep 173  
 QY 83 LKKMPKLTLLAIGMNFSTOKFTDMVATANNROTFVNSAIRFLRKXSPDGLDWEYR 142  
 DB 174 -adatqgnmlllqavrseldsyaay-akg-hhflislaapagpnykklfa-e-lg 228  
 QY 143 GSQSPAVDKEKERTTLVQDLANAFQOEAOQTSCKERLLLSAAVPAG-QTYVDAGLEVQKIA 201  
 DB 229 kvldylnmaydygswsnylghndanllyanpqpnatp-yntddavagylngvpankiv 287

```

OY 202 QUNLDFVNLMAYPDEHSGMEKVTGHNSPLYKROESSGAASLNDVAAYQOMLQKSTPASKLI 261
DB 288 lgmpllygrsfq-qt--eg-igkpynglgs-gsw--englwdykalpk-agatvkcdttak 339
OY 262 LGMPTGRSFTTLAASSDTRVGAPATSGTGPFTKEGMLAYEVCWKGATKORIODOK 321
DB 340 gcygydpsckelsfstdpamistekswlkqkgytunfw 378
OY 322 VPYIF-RD-NOMVGFDVDESFKTKVSYLKOKGLGGMVW 358

RESULT 6
ID CH1A.TRIHA STANDARD; PRT; 423 AA.
AC P48827;
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE 42 KD ENDOCHITINASE PRECURSOR (EC 3.2.1.14).
GN CHIT42.
OS TRICHODERMA HARZIANUM.
OC EUKARYOTA; FUNGI; DEUTEROMYCOTINA (IMPERFECT FUNGI).
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 35-52; 93-107; 371-385 & 397-414.
RX GARCIA I., LORA J.M., LA CRUZ J., BENITEZ T., LOBELLE A.,
RA PINOIR-TORO J.A.;
RL CURR. GENET. 27:83-89(1994).
CC -1- FUNCTION: MORPHOGENETIC ROLE DURING APICAL GROWTH/CELL DIVISION
CC AND DIFFERENTIATION (CELL WALL MORPHOGENESIS). ANTIFUNGAL
CC AGENT.
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE BETA-1,4 LINKAGES OF
CC N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- INDUCTION: SPECIFICALLY INDUCED BY CHITIN AND IS CATABOLITE
CC REPRESSED.
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC HYDROLASES).
DR EMBL; S78423; G999376; -.
KW HYDROLASE; GLYCOSIDASE; CHITIN DEGRADATION; SIGNAL; ZYMOGEN;
KM CHITIN-BINDING.
FT SIGNAL 1 22
FT PROPEP 23 34
FT CHAIN 35 423
FT ACT_SITE 171 171
FT CARBOHYD 218 218
FT SEQUENCE 423 AA; 46056 MW; 899DA50A CRC32;

Query Match 16.4%; Score 564; DB 2; Length 423;
Best Local Similarity 36.0%; Pred. No. 1,10e-99;
Matches 102; Conservative 70; Mismatches 94; Indels 17; Gaps 13;

DB 111 qltk-vtkarrglkvlslsgwtst-nfrsaastadnknfaktaitfkdvgfddi 168
OY 78 QENNGKAKKMPKKTLLAIGWNESTQKFTDMATANNROTFFVNSAIRFKRSFSDGLD 137
DB 169 dweyp-adatgaasmllllkevsgrdayaagy-apy-yhflitlaapagkdy-sklrl 224
OY 138 DWETPGSGSPAYDKERTTLVODLANAFQOEQOTSGKERLLLSAAPFAQOTVVDAGYEV 197
DB 225 adlqyldylnlmaydyagsfslptghdanlfnpsnnapc-fntdsavkdylngvpa 283
OY 198 DKLAQNLDVFNLMAYDEHSGMEKVTGHNSPLYKROESSGAASLNDVAAYQOMLQKSTPA 257
DB 284 nkliylgmpllygrsf--qntag--igqlyngvgs-gsw--eaglwkykalpk-agatvqd 335
OY 258 SKLILGMPTRGRSFTTLAASSDTRVGAPATSGTGPFTKEGMLAYEVCWKGATKORIOR 317
DB 336 svakgyysysakelsfstdpamistekswlkqkgytunfw 378
OY 318 ODOKVPIIFRD--NOMVGFDVDESFKTKVSYLKOKGLGGMVW 358

```

```

RESULT 7
ID CH1L.COCIM STANDARD; PRT; 427 AA.
AC P54196;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE ENDOCHITINASE 1 PRECURSOR (EC 3.2.1.14) (COMPLEMENT-FIXATION ANTIGEN)
DE (CF-ANTIGEN) (CF-AG).
GN CTSL.
OS COCCIDIOIDES IMMITIS.
OC EUKARYOTA; FUNGI; DEUTEROMYCOTINA (IMPERFECT FUNGI).
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C735;
RX MEDLINE; 96144270.
RA PISHKO E.J., KIRKLAND T.N., COLE G.T.;
RL GENE 167:173-177(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-SLVEIRA;
RA YANG C., ZHU Y., MAGEE D.M., COX R.A.;
RL SUBMITTED (MAR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE BETA-1,4 LINKAGES OF
CC N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC HYDROLASES).
DR EMBL; I41663; G1200190; -.
DR EMBL; U51271; G1286789; -.
DR EMBL; U33265; G1285728; -.
KW HYDROLASE; GLYCOSIDASE; CHITIN DEGRADATION; CHITIN-BINDING; SIGNAL;
KM GLYCOPROTEIN.
FT SIGNAL 1 2
FT CHAIN 2 427
FT CARBOHYD 387 387
FT CONFLICT 15 47
FT FT RWLSLRLCLFCELRGEMFTLSTVATVTTDIO -> VOAS
FT CONFLICT 199 199 K -> N (IN REF. 2).
SQ SEQUENCE 427 AA; 47629 MW; 6CB9AF73 CRC32;

Query Match 15.9%; Score 548; DB 2; Length 427;
Best Local Similarity 34.4%; Pred. No. 6,62e-96;
Matches 96; Conservative 70; Mismatches 94; Indels 19; Gaps 17;

DB 115 lkknnnlkllslsgwtst-nfkpaasteegrkfadtislknlkdjfgddidweyp 173
OY 83 LKKMNRKLTLLAIGWNESTQKFTDMATANNROTFFVNSAIRFKRSFSDGLDWEYP 142
DB 174 ede-kqandfvlllkacrealadaysak-hpngk-kflltlaapagpny-nk-1klaemd 228
OY 143 GSGSPAYDKERTTLVODLANAFQOEQOTSGKERLLLSAAPVAG-QTVVDAGYEVDKIA 201
DB 229 kyldfnlmaydfsgswkvsghmsnfvsttkp-estfcsdkavkdylykqvpankiv 287
OY 202 QUNLDFVNLMAYPDEHSGMEKVTGHNSPLYKROESSGAASLNDVAAYQOMLQKSTPASKLI 261
DB 288 lgmpllygrsf--ast-dg-igrsfngvg-gsw--engwlykdmqp-ggaqvteledia 339
OY 262 LGMPTGRSFTTLAASSDTRVGAPATSGTGPFTKEGMLAYEVCWKGATKORIODOK 321
DB 340 asygydknkryllysydtkvkaagkaeylckngmggm-w 377
OY 322 VPYIF-RDNOM-VGFDVDESFKTKVSYLKOKGLGGMVW 358

RESULT 8
ID CH1A.SERMA STANDARD; PRT; 563 AA.
AC P07254;
DT 01-APR-1988 (REL. 07, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE CHITINASE A PRECURSOR (EC 3.2.1.14).
GN CHIA.
OS SERRATIA MARCESCENS.

```

OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
OC ENTEROBACTERIACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA KOO J.C., LIM C.O., CHOI Y.J., KIM C.Y., BARK J.D., LEE S.Y.,  
RA CHO M.J.;  
RL SUBMITTED (JAN-1993) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 990 / QMB1466;  
RA JONES J.D.G., GRADY K.L., SUSLOW T.V., BEDBROOK J.R.;  
RL EMBO J. 5:467-473(1986).  
RN [3]  
RP REVISIONS, AND X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
RA MEDLINE: 95219379.  
RA PERAKIS A., TEWS I., DAUTER Z., OPPENHEIM A.B., CHET I., WILSON K.S.,  
RA VORIAS C.E.;  
RL STRUCTURE 2:1169-1180(1994).  
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE BETA-1,4 LINKAGES OF  
CC -1- N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.  
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
CC HYDROLASES).  
DR EMBL: L01455; G152818; -.  
DR EMBL: X03657; G46831; -.  
DR PIR: A25090; A25090.  
DR HSPD: P07254; 1CTN.  
DR PROSITE: PS01095; CHITINASE\_18.  
RW HYDROLASE; GLYCOSIDASE; CHITIN DEGRADATION; SIGNAL.  
FT SIGNAL 1 23  
FT CHAIN 24 563 CHITINASE A.  
FT DOMAIN 150 563 CATALYTIC.  
FT ACT\_SITE 315 563 PROTON DONOR (PROBABLE).  
FT ACT\_SITE 391 563 TA -> GP (IN REF. 2).  
FT ACT\_SITE 391 563 PROBABLE.  
FT CONFLICT 76 77 A -> P (IN REF. 2).  
FT CONFLICT 139 139 V -> I (IN REF. 2).  
FT CONFLICT 226 226 P -> A (IN REF. 1 AND 2).  
FT CONFLICT 395 395 PAMKPDATVTVNGVALLAQ -> RPSRRHLHGERROC  
FT CONFLICT 410 430 AAGQ (IN REF. 1 AND 2).  
FT CONFLICT 437 437 V -> I (IN REF. 2).  
FT CONFLICT 464 467 ATGP -> HRA (IN REF. 2).  
FT CONFLICT 473 473 K -> E (IN REF. 2).  
FT CONFLICT 484 484 G -> S (IN REF. 2).  
SQ SEQUENCE 563 AA; 60979 MW; 7764DB57 CRC32;  
Query Match 12.8%; Score 440; DB 2; Length 563;  
Best Local Similarity 29.1%; Pred. No. 1,04e-70;  
Matches 94; Conservative 83; Mismatches 123; Indels 23; Gaps 19;  
Db 231 pfaalqkqkytawddpykgnfgqlmaaklpahpdlkipsigvwtlsdpff-ngdkvk 289  
QY 57 AFAGMTNHOISTEWNDE-T-LYOEFGNGLKKNPKLTLAIGWNFSTQKFTDMVATN 114  
Db 290 -drftvsgvkeflgtwkdffgdvldwefpgykganpnlgspddgetyvlmlmelamldq 348  
QY 115 NKQTVNNAIRLKRKYSF-DGLDLDWEIPGSGS--PAV---DKERFTLVODLANAFQ 168  
Db 349 lsaetgr-kyeltsaisagkdkidkvvayvag--nsmdhiflmsydfygfalknlgpqt 405  
QY 169 EAQTSGERLRLLSAIVPAQGYVD-AGYEVDKIAQNLDFVNLMAVDFHGSME-KYTGHS 226  
Db 406 alnapawpdtayt-tvng-vnallagvkgpkkvvvgtvmgyrgvtygnygnipfgt 463  
QY 227 PLYKROESGAASLNVDAVQOVLQKTPASKLILGMPTGRSFTLASSDTRVGAAPAT 286  
Db 464 atg-pvkgvkwngivdyrgiaagfmgsewqlydataaepvfkpsetgdlifddarsvq 522  
QY 287 GSGTGPFFKKEGMLAYIVCS-WKQATQ-RIQD-QKVPYIFR-DN-QWVGFDVESRK 341  
Db 523 akgyvldkqlyglglfsweldadn 545  
QY 342 TKVSYLKQKGLGAGAMVWALDLD 364

RESULT 9  
ID CHIT\_NPVAC STANDARD; PRT; 551 AA.  
AC P41684;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
DE PROBABLE ENDOCHITINASE PRECURSOR (EC 3.2.1.14).  
OS AUTOGRAPH CALIFORNICA NUCLEAR POLYHEDROSIS VIRUS (ACMPV).  
OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; BACULOVIRIDAE; EUBACULOVIRINAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-C6;  
RX MEDLINE: 94303173.  
RA AYRES M.D., HOWARD S.C., KUZIO J., LOPEZ-FERBER M., POSSEE R.D.;  
RL Virology 202:586-605(1994).  
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE BETA-1,4 LINKAGES OF  
CC -1- N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.  
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
CC HYDROLASES).  
DR EMBL: L22858; G59195; -.  
DR PROSITE: PS00014; ER-TARGET.  
DR PROSITE: PS01095; CHITINASE\_18.  
RW HYDROLASE; GLYCOSIDASE; CHITIN DEGRADATION; SIGNAL; GLYCOPROTEIN.  
FT SIGNAL 1 17  
FT CHAIN 18 551 PROBABLE ENDOCHITINASE.  
FT ACT\_SITE 305 505 PROTON DONOR (BY SIMILARITY).  
FT CARBOHYD 173 173 POTENTIAL.  
FT CARBOHYD 444 444 POTENTIAL.  
FT SITE 548 551 PREVENT SECRETION FROM ER (POTENTIAL).  
SQ SEQUENCE 551 AA; 61368 MW; BDB124DB CRC32;  
Query Match 11.6%; Score 398; DB 2; Length 551;  
Best Local Similarity 29.5%; Pred. No. 4,28e-61;  
Matches 96; Conservative 79; Mismatches 122; Indels 28; Gaps 27;  
Db 221 pwaavqkbpqkyvsawn-ep-ykgnfgqlmaaklpahpdlkipsigvwtlsdp-fymhdv 277  
QY 57 AFAGMTNHOISTEWNDELYO-ENGLK--KM-NPKLTLAIGWNFSTQKFTDMVAT 112  
Db 278 -ekrnfvdsvkeflgtwkdffgdvldwefpgykganpnlgspddgetyvlmlmelamldq 336  
QY 113 ANNKOTEFVNSAIRFLKRKYSF-DGLDLDWEIPGSGS--PAV-D-K-E-R-FTLVODLANA 165  
Db 337 lddleagt-grv-yeltsaisagykliavvnya-e-agksjgkflmsydfkgaasndcl 392  
QY 166 FQO-EAQTSGKERLRLLSAIVPAQ-QTYVDAGYEVDKIAQNLDFVNLMAVDFHGSMEKYT- 222  
Db 333 gyqtvtva-ps-wmseelythvavdalikgvdnpkllivgvamyrgvtygnyndny 450  
QY 223 GHSNPLYKROESGAASLNVDAVQOVLQKTPASKLILGMPTGRSFTLASSDTRV 282  
Db 451 fsgtng-psgwtgdvdyrgldkldnnyyrtldsaagsyvdtkgldlfsdvs 509  
QY 283 APATSGTGPFFKKEGMLAYIVCSWKQATQRIQD-QKVPYIF-RDN-QWVGFDVES 339  
Db 510 vlgkvkyvdrnklyglfsweldadn 534  
QY 340 FTKVSYLKQKGLGAGAMVWALDLD 364  
RESULT 10  
ID CHIA\_ALTSO STANDARD; PRT; 820 AA.  
AC P32823;  
DT 01-OCT-1993 (REL. 27, CREATED)  
DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)  
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
DE CHITINASE A PRECURSOR (EC 3.2.1.14) (CHI-A).  
GN CHIA.  
OS ALTEROMONAS SP. (STRAIN O-7).  
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
OC VIBRONACEAE.  
RN [1]



Db 425 gsdgqk-ldaa-dygeaskyldwvnmtydffgawak-ngtaphaspilraygqigqfn 481  
Qy 184 VPAGQTYVADAGEVDKIAQNLDVFNLMADYFGHSMKVTGHNSPYKROEESGAAL-SLN 242  
Db 482 tadamakfkxkypackklllgfygrygtc-gvtgsap-gglatgpat-gty--eagled 536  
Qy 243 VDAAYQOMLQKGTSPASKILGMPYGRSFTLASSSDTRVGAPATGSGTGPFTKESGMALA 302  
Db 537 y-kvlnscapactgltag--tayahcgswsnydtpatlkshmdwaegqglgafwefsg 593  
Qy 303 YVEVCSMKGATKORIQDQKVPYIFRDQWVGFDVESFKTKVSYLKQKGLGAGAMWALDL 362  
Db 594 dtang 598  
Qy 363 DDFAG 367

RESULT 13  
ID CHIT\_STRPL STANDARD; PRT: 610 AA.  
AC P11220;  
DT 01-JUL-1989 (REL. 11, CREATED)  
DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)  
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
DE CHITINASE 63 PRECURSOR (EC 3.2.1.14).  
CN CHTA  
OS STREPTOMYCES PLICATUS.  
OC PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; STREPTOMYCETACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 92192480.  
RA ROBBINS P.W., OVERBYE K., ALBRIGHT C., BENFIELD B., PERO J.;  
RL GENE 111:69-76(1992).  
RN [2]  
RP SEQUENCE OF 1-45 FROM N.A., AND SEQUENCE OF 31-45.  
RX MEDLINE: 88087127.  
RA ROBBINS P.W., ALBRIGHT C., BENFIELD B.;  
RL J. BIOL. CHEM. 263:443-447(1988)  
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE BETA-1,4 LINKAGES OF  
N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.  
CC -1- INDUCTION: BY CHITIN.  
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
HYDROLASES).  
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.  
CC -1- SIMILARITY: CONTAINS A BACTERIAL-TYPE CELLOLOSE-BINDING DOMAIN  
(CBD).  
DR EMBL: M82804; G153216; -;  
DR EMBL: M18397; G153209; -;  
DR PIR: A29912; A29912.  
DR PIR: JH0573; JH0573.  
DR HSSP: P07254; ICTN.  
DR PROSITE: PS00018; EF\_HAND.  
DR PROSITE: PS00561; CBD\_BACTERIAL.  
DR PROSITE: PS01095; CHITINASE\_18.  
KV HYDROLASE; GLYCOSIDASE; CHITIN DEGRADATION; CHITIN-BINDING; SIGNAL.  
FT SIGNAL 1 30  
FT CHAIN 31 610 CHITINASE 63.  
FT DOMAIN 35 140 CELULOSE-BINDING.  
FT DOMAIN 148 229 FIBRONECTIN TYPE-III.  
FT DOMAIN 236 610 CATALYTIC.  
FT ACT\_SITE 383 383 PROTON DONOR (BY SIMILARITY).  
FT CONFLICT 3 3 F->I (IN REF. 2).  
SQ SEQUENCE 610 AA: 63974 MW: 255E8E35 CRC32;

Query Match 8.4%; Score 290; DB 2; Length 610;  
Best Local Similarity 26.7%; Pred. No. 4,24e-37;  
Matches 79; Conservative 62; Mismatches 134; Indels 21; Gaps 16;

Db 320 fnglnlkaeyphklllyfgwtsvsg-fpdavxnpaafakschdlvedprwadvfqgi 378  
Qy 77 YQEFNGLKMKNPRLKTLIAIGWMNFSTOKFTDMVATANNRQGFVNSAIFLKRYS-FDGL 135  
Db 379 dldwepnacyglscdetsapnafsmmkamreafgqdyilta-a---vtaadsgdqgk-id 433

Qy 136 DLDWEYPPSSQG-S--PAVDKERFTTLVQDLIANAFQEAQOTSGKERLLLSAANPAQTYVD 192  
Db 434 aa-dygeaskyldwvnmtydffgawak-ngtaphaspilnaydgiipgqgftadamakfk 491  
Qy 193 AGEYVDKIAQNLDVFNLMADYFGHSMKVTGHNSPYKROEESGAAL-SLVNDAAYQOML 251  
Db 432 skgypackklllgfygrygtc-gvtgsap-gglatgpa-gty--eagledy-kvlnsc 545  
Qy 252 QKGTSPASKILGMPYGRSFTLASSSDTRVGAPATGSGTGPFTKESGMALAYEVCSSMKG 311  
Db 546 patgvaq--tayahcgswsnydtpatlkshmdwaegqglgafwefsgdtng 599  
Qy 312 ATKORIQDQKVPYIFRDQWVGFDVESFKTKVSYLKQKGLGAGAMWALDDDFAG 367

RESULT 14  
ID KTYA\_KLULA STANDARD; PRT: 1146 AA.  
AC P09805;  
DT 01-MAR-1989 (REL. 10, CREATED)  
DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)  
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
DE KILLER TOXIN ALPHA AND BETA SUBUNITS PRECURSOR (REF2 PROTEIN)  
DE KENDUCHITINASE (EC 3.2.1.14).  
OS KLUYVEROMYCES LACTIS (YEAST).  
OC PLASMID pKRL-1.  
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CBS 2359/152;  
RA SOR F., FUKUHARA H.;  
RL CURR. GENET. 9:147-155(1985).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 84297209.  
RA STARK M.J.R., MILEHAM A.J., ROMANOS M.A., BOYD A.;  
RL NUCLEIC ACIDS RES. 12:6011-6030(1984).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 85037931.  
RA TOKUNAGA M., KAWAMURA A., HISHINUMA F.;  
RL NUCLEIC ACIDS RES. 12:7581-7597(1984).  
RN [4]  
RP IDENTIFICATION OF PROTEIN, AND SEQUENCE OF 30-44 AND 895-916.  
RX MEDLINE: 87004569.  
RA STARK M.J.R., BOYD A.;  
RL EMBO J. 5:1995-2002(1986).  
RN [5]  
RP SIMILARITY TO CHITINASE OF ALPHA-SUBUNIT.  
RX MEDLINE: 90259069.  
RA BRADSHAW H.D. JR.;  
RL NATURE 345:299-299(1990).  
RN [6]  
RP CHITINASE ACTIVITY OF ALPHA-SUBUNIT.  
RX MEDLINE: 91301161.  
RA BUTLER A.R., O'DONNELL R.W., MARTIN V.J., GOODAY G.W., STARK M.J.R.;  
RL EUR. J. BIOCHEM. 199:463-468(1991).  
CC -1- FUNCTION: THE ALPHA SUBUNIT IS A POTENT EXOCHITINASE, ALONG WITH  
THE BETA SUBUNIT IT PLAYS A ROLE IN THE INITIAL INTERACTION OF THE  
TOXIN WITH SENSITIVE CELLS AND ALLOW THE GAMMA SUBUNIT (THE ACTIVE  
TOXIN) TO GAIN ENTRY INTO THE CELL.  
CC -1- P1M: REF2 IS POTENTIALLY SPLIT BY MEMBRANE-BOUND BASIC AMINO ACID-  
SPECIFIC PEPTIDASE TO YIELD THE ALPHA AND BETA SUBUNITS.  
CC -1- SUBUNIT: THE KILLER TOXIN IS COMPOSED OF THREE SUBUNITS: ALPHA,  
BETA AND GAMMA.  
CC -1- SIMILARITY: LOCAL, TO OTHER CHITIN-BINDING PROTEINS.  
CC -1- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.  
DR EMBL: X07127; G2829; -;  
DR EMBL: X00762; G2844; -;  
DR EMBL: X01095; G2849; -;  
DR PIR: S07915; S07915.  
DR HSSP: P02877; 1HEV.  
DR PROSITE: PS00026; CHITIN\_BINDING.







Db	200	riififisvldhngsvskvrgnhsplfs-l-pgdbks-sa-yammywvjaypvekl	1254
Qy	181	QNDFVNLMAIDFNGSEKTKTGHNSPLKROEESGAASLWDAVQOMLQKGPASKLI	240
Db	255	mg1btygrtfhllkasqnelragavspaspkytkdgafllayelccfvrrakkrwinq	314
Qy	241	LGMPTYGRSFTLASSSDTRVGCAPATGSGTDPFPRKEGMALAYVVCWM-KGATQRQRDQ	299
Db	315	yvpafafgkewvgvddaisfygkafllkrehfgamwvclldldfryfcgtpfpflvht	374
Qy	300	KVPPIFRNOMVGGDDVESFKTSYLKOKGLGSGAMWALDLDDFAFGSCNOGRYPPIQT	359
Db	375	lnnll	379
Qy	360	LROEL	364

ID	RESULT
AC	R73993 standard; protein; 666 AA.
DT	10-JAN-1996 (first entry)
DE	Hamster oviduct specific glycoprotein.
KM	Hamster oviduct specific glycoprotein; recombinant production;
OS	HOOP.
FT	Criceululus griseus.
FT	Key
FT	Peptide
FT	/label= sig_peptide
FT	Peptide
FT	/label= mat_peptide
PN	U07107979-A.
PD	25-APR-1995.
PF	15-AUG-1994; 214227.
PR	19-AUG-1993; JP-227881.
PA	(KINO-) KINOSEI PEPTIDE KENKYUSHO KK.
DR	WPI: 95-190179/25.
DR	N-PSDB; Q90444.
PT	New DNA encoding an oviduct-specific glycoprotein - useful for
PT	recombinant protein production in high quantities.
PS	Claim 4; Pages 16-18; 22pp; Japanese.
CC	G90444 encodes R73993 hamster oviduct specific glycoprotein (HOOP)
CC	The cDNA can be used for the commercial recombinant prodn. of
CC	HOOP in high quantities.
Q	Sequence 666 AA;

Query Match	47.9%;	Score 1303;	DB 14;	Length 668;
Best Local Similarity	48.4%;	Pred. No. 2.08e-111;		
Matches 177;	Conservative	81;	Mismatches 99;	Indels 9;
				Gaps 7;

[illegible]

```

Db      374  innell  379
      | : |
QY      359  TLRQEL  364

```

ID	RESULT	3
AC	R73992; standard; Protein; 718 AA.	
DT	R73992;	
DE	10-UMN-1996 (first entry)	
KW	Murine oviduct specific glycoprotein.	
KW	Murine oviduct specific glycoprotein; recombinant production;	
OS	Mus musculus.	
FT	key	Location/Qualifiers
FT	peptide	1..18
FT	/label= sig_peptide	
FT	peptide	19..718
PN	/label= mat_peptide	
PD	25-APR-1995.	
PF	15-AUG-1994; 214227.	
PR	15-AUG-1993; JP-227681.	
PA	(KINO-) KINOSEI PEPRIDE KENKYUSHO KK.	
DR	WPI: 95-190179/25.	
DR	N-PSDB: Q90443.	
PT	New DNA encoding an oviduct-specific glycoprotein - useful for	
PT	recombinant protein production in high quantities.	
PS	Claim 4; Pages 11-14; 22p; Japanese.	
CC	G90443 encodes R73992 murine oviduct specific glycoprotein (MOGP);	
CC	The cDNA can be used for the commercial recombinant prodn. of	
CC	MOGP in high quantities.	
CC	Sequence 718 AA;	
CC		

Query Match	47.9%;	Score 1303;	DB 14;	Length 718;
Best Local Similarity	49.9%;	Pred. No. 2,08e-111;		
Matches 182;	Conservative	76;	Mismatches 100;	Indels 7;
				Gaps 4;

[illegible]

PD 30-JAN-1988.  
 PF 16-JUL-1986; 167518.  
 PR 16-JUL-1986; JP-167518.  
 PA (DAIN) Dainippon Pharm KK.  
 DR WPI: 88-068419/10.  
 P-PSDB: PB1342.  
 PR New polypeptide and DNA encoding it - related to protective  
 PT mechanisms such as immune response etc.  
 PS Disclosure: P: Japanese.  
 CC This polypeptide is involved in protective mechanisms such as immune  
 CC response, cell growth and activation of protective functions.  
 SQ Sequence 321 AA;

Query Match	32.4%;	Score 881;	DB 1;	Length 321;
Best Local Similarity	50.2%;	Pred. No. 9.90e-71;		
Matches	108;	Conservative	60;	Mismatches 41;
			Indels	6;
			Gaps	3;

```
Db      22 klvcyvtswsgyregdscfpdaldrflcthllysfaaisndhidtwewndvltgym|nt 81
        ||||:::||::: | | ||||::|| ::: | ||| || : |
QY     2 klvcfETNMAQYRQGAEARFLPKDDPSLCTHLIYAAGMTNNQLSTENNDFTLYQEFNG 61
```

```
Db      82 lknnpnlkclsvvgwvngfsqrfsklasntqtsrrtflksvpfflithgfdvdlawlyp 141
        ||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Qy     62 LKKANPKLKTLLAIGGWNFGTQKFTDWATANNRQT FVNSAIRFLAKYSPFDGLDDEYF 122
```

```
Db 142 g*rg----dqhfttlllkemkaefikeaq-pgkxqlllsalsagkvrtidssydiarlsq 195
   | :| :|:|:|:|:| :| :| :| :|:|:|:|:| :|:|:|:|:|
QY 122 GSGSPAVDKERFTTLVODLANAFQEQETSGKERLLLSA VPAGQTYVDAGYEVDKIAQ 184
```

**D**b    196 hldlslmtydfnhgawrgtcqhsplfrqgedasp    230  
         :||:::|||::|    |||::|::|::|  
**Q**y    182 NLDFVNIMAYDHFHGSWEKVTGHSNPLKYKRQEESA    216

RESULT	5
ID	R70025 standard; Protein; 371 AA.
NC	P70025

DT 29-SEP-1995 (first entry)  
DE Tobacco chitinase encoded by tobacco Cluster-A cDNA clone ca-3  
KW Tobacco; chitinase; antifungal; fungicide; Cluster-A.  
OS Nicotiana glauca

	Location/Qualifiers
FH Key	1..18
FT Peptide	/label= signal peptide
FT /label=	signal peptide
ET Misc difference	75

ET	/note- "not present in genomic clone"
ET	Misc-difference 182..195
FT	/note- "Determined by protein sequencing"
ET	Misc difference 216..241

	Protein	Accession	RefSeq	Gene	Protein	Accession	RefSeq	Gene
EP	639642-A							
PD	22-FEB-1995							
PF	17-AUG-1993							

PA (MOGE-) MOGEN INT NV.  
PA (UYLE-) RIJKSUNIV LEIDEN.  
PI Apotheker-de Groot M., Bol JF, Cornelissen BJC, Linthorst HJM

DR WPI: 95-083454/12.  
DR N-PSDB: Q82973.  
PT New plant protein having endo-chitinase activity - used in

PT less susceptible to fungal infection.  
PS Claim 3; Page 19-20; 43pp; English.  
CC Screening of a lambda ZAP cDNA library of TMV-infected Samsun

CC CDNA clone) resulted in the isolated of 11 positively hybridising  
CC clones. Analysis revealed that all were identical and corresp. to  
CC Cluster-A CDNA. The nt sequence of clone CA-3 is given in Q82973/  
CC 370025.

CC cdna clone was completed by performing a PCR reaction on clone CA  
CC with primer T7 (Q82794) and primer P1 (Q82975). A partial AA  
CC sequence of the protein isolated from TMV-infected tobacco leaves

CC was obtd. and found to be almost identical to the AA sequence  
CC deduced from the cDNA.  
SQ Sequence 371 AA;

Query Match	21.98;	Score 597;	DB 13;	Length 371;
Best Local Similarity	33.08;	Pred. No. 9.86e-44;		
Matches 107; conservative	86;	Mismatches 108;	Indels 23;	Gaps 18;

Dd 36 mndstfthlfcadlnpsngllispengds-fsqgfstvgqrknpvsktfsiaaguu  
:::| | | | : : | : : : : : : : : : : | | : | | :  
Qy 23 KDLDPSTCTHLIYAFAGMT--NHQLSTTEWNDETLLYQEENG-LKKMNPKLKTLIAI-GG

Dd 95 a-dttagymarqpnsrksfsslsrlarqfyngldlweyp--l-s-atcmtnlglill 149  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 79 nfgioKFTDMATANNRQTfVNSAIRFLRKYSFSGLDLMEWYPGSGSSPAVDKERFTTLIV 138

D5 newrcainmeaingsgraaillcvausysprvngjnypvcsaarunwlinmgydtygnw 20  
 : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 QY 139 QDLANFQOEQTSGKERLLSAVPAGOTYYDAGYEVDKTAQNLDVFNLMAYDFHGS-W 19

[illegible]

204    **inattapaagsnvga**-**vaagsh**-**lynnriaylvqr**atraty<sup>natlvga</sup>cy<sup>gshwisy</sup> 32  
       : ||| : | : : | : | : : | : : | : : | : : :  
 QY    257 **DTRVGA**.**PATGSGT****PGPFTKEGSH**LA<sup>YEVCSWKGATQR</sup>-**IQDQ**-**VP**-**YIFRD**NQWGF 31

342 ucqsvluuvyvgigligiyaw 34  
314 DDVESFKTKVSYLKQKGLGAMVW 337

RESULT	6
ID	R70029 standard; Protein; 377 AA.
AC	R70029;

DE Tobacco cluster-A protein encoded by genomic clone.  
KW Tobacco; chitinase; antifungal; fungicide; Cluster-A  
OS *Nicotiana tabacum*.

FT	peptide	1..25
FT	/label= signal peptide	
FT	Misc_difference	28..29

FT	Misc_difference	188..201
FT	/label=	Determined by sequencing the protein
FT	Misc_difference	222..247

PN	EP-639642-A.
PD	22-FEB-1995.
PF	17-AUG-1993; 202425.
17	AUG 1993. ED 200405

PA (MOGE-) MOGEN INT NV.  
PA (UYLE-) RIJKSUNIV LEIDEN.  
PI Apotheker-de Groot M, Bol JF, Cornelissen BJC, Linthorst HJM,  
Pi Melchers RF, Deontia AS, Gelhaarwaase M.

DR WPI; 95-083454/12.  
DR N-PSDB; Q82976.  
PT New plant protein having endo-chitinase activity - used in  
PT antifungal compounds and to develop transformed plants which are

PI less susceptible to fungal infection.  
PS Claim 3; Page 25-28; 43pp; English.  
CC Screening of a lambda ZAP cDNA library of TMV-infected Samsun NN  
CC tobacco plants with a probe derived from PROR40 (a partial Cluster-2

CC clones. Analysis revealed that all were identical and corresponded to the sequence of cDNA clone CA-3 given in Q82973/  
CC R70025. A genomic library of *N. tabacum* was screened using the  
CC cDNA clone) resulted in the isolation of 11 positively hybridizing

CC The complete nt. sequence of Cluster-A cDNA including the deduced  
CC primary structure of the Cluster-A protein the 5' and 3' UTR regions  
CC of the gene are shown in 082976/R70029. Comparison of the cDNA  
CC cluster-A cDNA insert of clone CA-3 as a probe (see 082976, 082978).







QY 206 PLVKRQESGAAASLNDAAVQOVLKGTGPAKLLIGMPTYGSGFTLASSSDIRVGPAT 265  
 DB 463 hravkgtw- englvdytrqiasqfmsgewgyltdgteapvyfkpsgtldltdsdarsvg 520  
 QY 266 GSOTPGFTKEGMLAYVECS-WKCATKQ-RIDQ-KVYIFR-DN-QWVGFDVSEFK 320  
 DB 521 akgyvldkqlgglfsweldadn 543  
 QY 321 TKVSYLKQKGLGGAMWALDLD 343

RESULT 14  
 ID R32547 standard; Protein: 238 AA.  
 AC R32547;  
 DT 17-JUN-1993 (first entry)  
 DE HanPV IE-1 gene protein (partial sequence).  
 KW Heliothis armigera nuclear polyhedrosis virus; Ha; NPV;  
 KM insecticide; immediate early-1.  
 OS Heliothis armigera nuclear polyhedrosis virus.  
 PN MO9303144-A.  
 PD 18-FEB-1993.  
 PF 05-AUG-1992; AC0413.  
 PR 05-AUG-1991; AU-007576.  
 PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
 PI Christian PD, Dall DJ, Gordon KHJ, Hanzlik TN, Srisanthan A.  
 DR WPI: 93-076501/09.  
 N-PSDB: Q35982.  
 PT Insect virus with reduced capacity to occlude viral particles - used  
 PT for controlling proliferation of insect pests without horizontal  
 PT transmission  
 PS Disclosure: Fig 2: 51pp; English.  
 CC The sequence is that encoded by the Heliothis armigera (Ha)  
 CC nuclear polyhedrosis virus (NPV) immediate early-1 (IE-1) gene  
 CC (partial sequence).  
 SQ Sequence 238 AA;

Query Match 10.0%; Score 271; DB 6; Length 238;  
 Best Local Similarity 34.1%; Pred. No. 9.19e-14;  
 Matches 58; Conservative 43; Mismatches 51; Indels 18; Gaps 14;

DB 44 erdkflsnpwrefvltwfk-fdgididwefpgskganpvgdverdnmtyaligela 102  
 QY 94 NRQTFVNSAI-R-F-L-RKISFDGLDMDWETPESQGS-PAV-DKER---FTLVODLAN 143  
 DB 103 mldq-vqigturtlettsasgldkaav-ncdraqyldkifmsydfkgawntdlg 160  
 QY 144 AFQGEAQTSGKERLLLSAAVPAGQTYVDAGYEVDKIAQNLDYVNLMAYPDHFSGMEKVT-G 202  
 DB 161 hgtaly--gsaawkpnepytanvavdallagrvnp-kklvlgvamygrgwt 207  
 QY 203 HNSPLYKROESGAAASLNDAAVQOVL-OKGTGPAKLLIGMPTYGSGFT 251

RESULT 15  
 ID W02156 standard; Protein: 1046 AA.  
 AC W02156;  
 DT 14-JAN-1997 (first entry)  
 DE periplasmic chitodextrinase.  
 KW periplasmic chitodextrinase; periplasmic Beta-N-acetylglicosaminidase;  
 KW Beta-N-acetylglicosaminidase; chitin; oligosaccharide; catabolic;  
 KW catabolism.  
 OS Vibrio furnissii.  
 PN W09625424-A1.  
 PD 22-AUG-1996.  
 PF 13-FEB-1996; U02332.  
 PR 13-FEB-1995; US-386727.  
 PA (UXTD ) UNIV JOHNS HOPKINS.  
 PI Baessler B, Chittiaru E, Keyhani N, Roseman S, Rowe C;  
 PI Yu C;  
 DR WPI: 96-393335/39.  
 DR N-PSDB: T36387.  
 PT Chitin biosynthetic enzymes end I, exo I and exo II - are

PT periplasmic chito:dextrinase(s), periplasmic beta-Glucanidase(s) and  
 PT aryl beta-N-acetyl:gluco:amidase(s), respectively  
 PS Claim 2; Page 68-71; 101pp; English.  
 CC Periplasmic chitodextrinase (W02156), periplasmic  
 CC beta-N-acetylglicosaminidase (W02157) and aryl  
 CC beta-N-acetylglicosaminidase (W02158) can be used to produce chitin  
 CC oligosaccharides with the structure (GlcNAc)n where n is 2 or  
 CC higher, by contacting them with soluble chitin. The enzymes are  
 CC encoded by the genes endI, exoI and exoII respectively. They are  
 CC all genes involved in the catabolic pathway of chitin.  
 SQ Sequence 1046 AA;

Query Match 8.0%; Score 219; DB 19; Length 1046;  
 Best Local Similarity 33.3%; Pred. No. 3.07e-09;  
 Matches 48; Conservative 40; Mismatches 42; Indels 14; Gaps 10;

DB 454 etfadaavemmkkyrfidgididlrinsidgltgnpddtfsesrraylmnsyhelmrvir 513  
 QY 96 QTFVNSAIRLRKYSFDGLDLD-W--EYPSQSPV-VP-KE-RFTTV---QDLNARQ 146  
 DB 514 ekldvasagdgvyhmltlaaps-ayllrgmetavtgyldvynlmsydlhgawndhvg 572  
 QY 147 QEAQ-TSGKE--RLLSAAVPAGQTYVDAGYEVDKIAQNLDYVNLMAYPDHFSGMEKVTG 203  
 DB 573 naalydtgkdselaq-nygttaq 595  
 QY 204 NSPLYKROESGAAASLNDAAVQ 227

Search completed: Fri Jun 27 17:04:25 1997  
 Job time : 74 secs.







```

OY 62 LKKNPKLTKLLAIGMNEFTQKFTDMVATANNRQTFVNSAIFLRKRSYDGLDMWEXP 121
DB 143 grr-----dkqhtllkemekaeifkeaq-pgkqklllssalsagvltlssydlaklsq 196
OY 122 GSGGSAVNDKERFTTLVODLANAFQOEAQTSKGERLLLSAAVPAQGTYYDAGYEVDKIAQ 181
DB 197 hldfslmtydfhgwargtqthsp1frgqdasdpdrfsmtdyavymrlrgapask1vm 256
OY 182 NUDFVNLMAVDFHGSWEKVTGHSNPLYKROESGAASLNVDAVQOMLOKGTIPASKLIL 241
DB 257 glptfgrsfllaas-etyvqapispigipgrfkeagtlayelcdflrgatvhrtlgqg 315
OY 242 GMPTYGRSFTLASSSDTRVGAPATGSGTGPFTKEGMLAYEVCWSV-KGATKORIQDOK 300
DB 316 vpyatgngvvygddgsvskvqylkdrqlaagmwaldlddfgscfgqdlrfrpltna 375
OY 301 VPIYIFRDNONWGFDDVESFKTKVSYLKOKGLGAMWALDLDDEFAGFSCNOG-RYPLIOT 359
DB 376 lkda1a 381
OY 360 LRQELS 365

RESULT 2
ENTRY 2
TITLE S51327 #type complete
ORGANISM heparin-binding glycoprotein 38k - pig
#formal_name Sus scrofa domestica #common_name domestic pig
#accession 07-May-1995 #sequence_revision 07-May-1995 #text_change
#authors S51327
#submission Shackson, L.M.; Mann, D.M.; Millis, A.J.T.
#description Submitted to the EMBL Data Library, January 1995
#description Identification of a 38kDa heparin-binding glycoprotein
(gp38k) in differentiating vascular smooth muscle cells as
a member of a group of proteins associated with tissue
remodeling.
#accession S51327
#status Preliminary
#residues 1-383 #label SHA
#cross-references EMBL:247803
SUMMARY #length 383 #molecular_weight 42443 #checksum 2907

Query Match 52.5%; Score 1428; DB 16; Length 383;
Best Local Similarity 52.2%; Pred. No. 2.76e-239;
Matches 191; Conservative 81; Mismatches 85; Indels 9; Gaps 5;

DB 23 klvcytswsgyregdscfpda1pflcthllyfamsnne1dltewndvlydltnt 82
OY 2 klvcyftnmaoyrgearflpKDLPDLCTHLIYAFAGMTNHQLSTENMDETLYQEFNG 61
DB 83 klkrrnplkllsvgwmfsgqrfsklaasntqgrtfkssvppflrtfthgfdgldawisp 142
OY 62 LKKNPKLTKLLAIGMNEFTQKFTDMVATANNRQTFVNSAIFLRKRSYDGLDMWEXP 121
DB 143 grr-----dkrhtllvkemekaeifvreal-pqterlllsgvasaokvaidrgydlqisq 196
OY 122 GSGGSAVNDKERFTTLVODLANAFQOEAQTSKGERLLLSAAVPAQGTYYDAGYEVDKIAQ 181
DB 197 hldfslmtydfhgwargtqthsp1frgqdasdpdrfsmtdyavymrlrgapask1vm 256
OY 182 NUDFVNLMAVDFHGSWEKVTGHSNPLYKROESGAASLNVDAVQOMLOKGTIPASKLIL 241
DB 257 glptfgrsfllaas-tdvapasgpgipgrfkeagtlayelcdflrgatvhrtlgqg 315
OY 242 GMPTYGRSFTLASSSDTRVGAPATGSGTGPFTKEGMLAYEVCWSV-KGATKORIQDOK 300
DB 316 vpyatgngvvygddgsvskvqylkdrqlaagmwaldlddfgscfgqdlrfrpltna 375
OY 301 VPIYIFRDNONWGFDDVESFKTKVSYLKOKGLGAMWALDLDDEFAGFSCNOG-RYPLIOT 359
DB 376 lkda1a 381
```

```

OY 360 LRQELS 365

RESULT 3
ENTRY 3
TITLE S61550 #type complete
ORGANISM BRP39 protein - mouse
#formal_name Mus musculus #common_name house mouse
#accession 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change
#authors S61550
#submission Morrison, B.W.
#description Submitted to the EMBL Data Library, November 1995
#accession S61550
#status Preliminary
#residues 1-381 #label MOR
#cross-references EMBL:X93035
SUMMARY #length 381 #molecular_weight 43001 #checksum 5107

Query Match 52.1%; Score 1417; DB 14; Length 381;
Best Local Similarity 50.7%; Pred. No. 3.51e-237;
Matches 186; Conservative 84; Mismatches 86; Indels 11; Gaps 9;

DB 23 klvcytswsgyregdscfpda1pflcthllyfamsnne1dltewndvlydltnt 82
OY 2 klvcyftnmaoyrgearflpKDLPDLCTHLIYAFAGMT-NHQLSTENMDETLYQEFN 60
DB 83 klkrrnplkllsvgwmfsgqrfsklaasntqgrtfkssvppflrtfthgfdgldawisp 142
OY 61 GKKMNPRLKTLAIGGMVFGTQKFTDMVATANNRQTFVNSAIFLRKRSYDGLDMWEXP 120
DB 143 p--r--lr--dkgyfslklhelnaefkveq-pgreklllssalsagvltlssydlaklsq 196
OY 121 PGSGSAVNDKERFTTLVODLANAFQOEAQTSKGERLLLSAAVPAQGTYYDAGYEVDKIA 180
DB 197 hldfslmtydfhgwargtqthsp1frgqdasdpdrfsmtdyavymrlrgapask1 256
OY 181 QNIDFVNLMAYDFHGSWEKVTGHSNPLYKROESGAASLNVDAVQOMLOKGTIPASKLIL 240
DB 257 mgjptfsgfllaas-englgapispigipgrfkeagtlayelcdflrgatvhrtlgqg 315
OY 241 LGMPTYGRSFTLASSSDTRVGAPATGSGTGPFTKEGMLAYEVCWSV-KGATKORIQDOK 299
DB 316 kvpfatgngvygwhesvknkvglkxkklagammwaldlddfg-tcqpkeftpln 374
OY 300 KVPYIFRDNONWGFDDVESFKTKVSYLKOKGLGAMWALDLDDEFAGFSCNOGRY-PLIQ 358
DB 375 alkda1a 381
OY 359 TLQELS 365

RESULT 4
ENTRY 4
TITLE I48271 #type complete
ORGANISM BRP39 protein - mouse
#formal_name Mus musculus #common_name house mouse
#accession 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change
#authors I48271
#submission Morrison, B.W.; Leder, P.
#description Oncogene (1994) 9:3417-3426
#description neu and ras initiate murine mammary tumors that share genetic
markers generally absent in c-myc and int-2-initiated
tumors.
#cross-references MIMD:95060797
SUMMARY #accession I48271
#status Preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-381 #label RES
#cross-references EMBL:X93035; NID:g1085065; CDS_PID:g1085066
```



ENTRY	7	#type complete
146470	estrogen dependent oviduct protein precursor - sheep	
#formal_name	Ovis orientalis aries, Ovis ammon aries	
#common_name	domestic sheep	
16-Aug-1996	#sequence_revision	16-Aug-1996
16-Aug-1996	#text_change	
146470		
146470		
Desouza, M.M.; Murray, M.K.		
Endocrinology (1995) 136:2485-2496		
An estrogen-dependent secretory protein, which shares identity with chitnases, is expressed in a temporally and regionally specific manner in the sheep oviduct at the time of fertilization and embryo development.		
#cross-references	MUTID:95269691	
#accession	I46470	
#molecule_type	preliminary; translated from GB/EMBL/DBJ	
#status	preliminary	
#residues	1-539	#label DES
#cross-references	EMBL:U61719, NID:9885600, CDS_PID:9885601	
#length	539	#molecular_weight 59535 #checksum 2120
Query Match	48.4%; Score 1318; DB 14; Length 539;	
Best Local Similarity	50.1%; Pred. No. 2,92e-218;	
Matches	183; Conservative 77; Mismatches 98; Indels 7; Gaps 5;	
Db	23 klvlyfcnwfatsrpsasllprldldflctchlyfafasmmnglypkdpldekltypefn 82	
Qy	2 klvlyfnnwagyrqgehrflpkldpslchtllyafagmtlnolstye-wndetlyoefn 60	
Db	83 klkerngjkllvlygwmfgrtsrftkmlstfsmferfkvsiallrlthgfdqlfflyl 142	
Qy	61 glkkmprklkllalagmhfctokftdmvatannrotfvsairflrkksfsgldlmedy 120	
Db	143 pglgspardtwftvflleelqafkneagltmrrplllisaavsgdhvldqkaydarllg 202	
Qy	121 pgsogspavndekrrpttlyodlaanafoeaoqtsgkerllsaaavpaqotvvdagvevdkia 180	
Db	203 rllldffisvlysdlnsgwkyrtgmhnsplfsipgdpksa-yan--sy- wrqlgvppekl 257	
Qy	181 onldfwnlmaadefrgsmekvtghnsplyrkqessgaaslaandvaayoolqctgtpaskli 240	
Db	258 mglpctygrtfnlllaasneqlagagaapspkytkbaqfllayvevcsfygrakkrwindq 317	
Qy	241 lgmfttygrstfllassdptrrgarpatgctgtptrkcgmlatayevcs-kgatkrldo 299	
Db	318 yvypafkfygkwvygdaisfykafikrehfygamwrlldldfngnfcgtpfplah 377	
Qy	300 kvprifrdnqmvghpdivesfkrvksylkqglgagawmaldddfagfscnqgrvplq 359	
Db	378 lnnll 382	
Qy	360 lroel 364	
RESULT	8	
ENTRY	A53918	#type complete
TITLE	chitinase (EC 3.2.1.14) precursor - braconid wasp (Chelonus sp.)	
ORGANISM	#formal_name Chelonus sp.	
DATE	28-Jul-1995	#sequence_revision 28-Jul-1995
	03-May-1996	#text_change
ACCESSIONS	A53918	
REFERENCE	A53918	
#authors	Krishnan, A.; Nair, P.N.; Jones, D.	
#journal	J. Biol. Chem. (1994) 269:20971-20976	
#title	Isolation, cloning, and characterization of new chitinase stored in active form in chitin-lined venom reservoir.	
#accession	A53918	
#status	preliminary	
#molecule_type	mRNA	

##residues	1-483	##label	KRI
##cross-references	GB:U10422		
KEYWORDS	glycosidases; hydrolase		
SUMMARY	#length 483 #molecular-weight 52013 #checksum 5938		
Query Match	43.3%;	Score 1179;	DB 12; Length 483;
Best Local Similarity	42.5%;	Pred. No. 8,78e-192;	
Matches	158; Conservative	98; Mismatches	99; Indels 17; Gaps 11;
Db	23	kvcvfygawsvyrgngkfndngidpclclhlysfvngvdkvxdlpdwsdipnldgt	82
OY	2	KLVCYFTWMAOYRGGEARFLPKDLPCLCHLYAFAGMNHOLSTTE-WND--ETL--Y	56
Db	83	gkfcsllrkhpvsyrlmawvgmagsvpfsgmasdqtatreaqnvvklfgyqgfdgd	142
OY	57	QEENGLKMNKRLTLTLAGWMNGTGFEDWATANNROTFFNSARFLFRKRSYFSGDLT	116
Db	143	dweypaaggsapadvkmvklckalkkafvg--h----d-yllsaavaapetsasxeyd	195
OY	117	DWEYFGSGSGSAVDKERFTTLYVDLANAFQEOAOTSEKREKLILISAAPACQTYVDAGYEV	176
Db	196	aemsqyldflnltmydfihpwpdgtctgmhappssashdsngelnkvnkaavkywldngvpk	255
OY	177	DKIAQNLDEVMALMYDEPHGSWEKTYGNSPLPKRQESGAASLNDVAAYQWMLQKSTPA	236
Db	256	eklvavgpavagksftlinspnkglgavpsgaagtagpytgegnlllyneicemgkagdwv	315
OY	237	SKLLIGMPTYRSFTLLASSDTRGAAATSGTPGPTKSGMLAYEYCSW-KGATKQR	295
Db	316	vqdhckgyrvavkngwvsfddlaakagkfgllqeglgagmwsletddfkgl-ge-k	373
OY	296	IQD-QK-VPIYFRNOMVGDDVESEFKVSYLKQKGLGAMWALDDDFAGSCNQGR	353
Db	374	ypvklakalsvlg	385
OY	354	YPLIQTLRQELS	365
RESULT	9		
ENTRY	A38221	#type complete	
TITLE	chitinase (EC 3.2.1.14) MFI - nematode (Brugia malayi)		
ORGANISM	#formal_name Brugia malayi		
DATE	04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 03-May-1996		
ACCESSIONS	A38221		
REFERENCE	A38221		
authors	Fuhrman, J.A.; Lane, W.S.; Smith, R.F.; Plessens, W.F.; Perlner, F.B.		
#journal	Proc. Natl. Acad. Sci. U.S.A. (1992) 89:1548-1552		
#title	Translucence blocking antibodies recognize microfilarial chitinase in Brugian lymphatic filariasis.		
#cross-references	MUID:92179220		
#accession	A38221		
##status	preliminary; not compared with conceptual translation		
##molecule_type	nucleic acid; protein		
##residues	1-504 #label FWH		
##cross-references	NCBIP:85345		
##note	sequence extracted from NCBI backbone		
KEYWORDS	glycosidases; hydrolase		
SUMMARY	#length 504 #molecular-weight 55971 #checksum 7343		
Query Match	39.9%;	Score 1086;	DB 12; Length 504;
Best Local Similarity	42.8%;	Pred. No. 3,90e-174;	
Matches	158; Conservative	94; Mismatches	102; Indels 15; Gaps 11;
Db	27	cyntinwagrydgegkflpnpnlnglchillyafakvdeIdgskpfewndedtwskmys	86
OY	5	CYFNWAOYRGGEARFLPKDLPCLCHLYAFAGMNHOLSTT-EWND-E-T-----LYQ	57
Db	87	avtklirenpjklvllsygygnfsaiftgljakagqterfifsasalfiknmfdgfdld	146
OY	58	EFNGLKRNPKRLKTLTLAGWMNGTGFEDWATANNROTFFNSARFLFRKRSYFSGDLT	117



```
##experimental source cv. Samsun NN  
#accession SS1632  
#molecule_type protein  
#residues 31-36;87-112;252-275;282-305;337-371 ##label HEW  
CLASSIFICATION #superfamily Streptomyces chitinase chl40  
KEYWORDS glycosidase; hydrolase  
FEATURE  
1-25 #domain signal sequence #status predicted #label SIG\  
26-371 #product chitinase/Lysozyme PZ #status predicted #label  
MAT  
371-378 #domain carboxyl-terminal propeptide #status  
experimental #label PRO  
SUMMARY #length 378 #molecular_weight 42019 #checksum 6689
```

```
Query Match 21.8%; Score 594; DB 5; Length 378;  
Best Local Similarity 32.7%; Pred. No. 2,86e+82;  
Matches 106; Conservative 87; Mismatches 108; Indels 23; Gaps 18;
```

```
D43 mndstlftlhfcadlnpsnqlispengds-fqsfstvgrkpsvxtflslagrr 101  
:::|::||::||::|::|::|::|::|::|::|::|::|::|::|::|::|:  
QY 23 KDLDSLTCHLIYAAGNT-NHOLSTEWDELHYLOEFNG-LKKMPKRLTLAI-GGW 78  
  
D102 a-dttaagimarqprrskfsidsrlarfgfnglddweyr--L-s-atdmnlgll 156  
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
QY 79 NFGTKRFDMATANNRRTFVNSAIRFKRKISFEGLDDIMEPSQGSPAVDKERFTLV 138  
  
D157 newtalmearnsgraalllaavsyprvrglpnyvesvarnlalnmaydfgpwv 216  
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
QY 139 QDLANAFQEQAQTSGKERLLLSAIVPAQGVTVDAGEVDKLAQNLDPEVNLMADVPHS-W 197  
  
D217 spsgfnhaqfldvpnh--lsqs---dg-lhawiqaypvtkklvlglpflyawrllvnp 270  
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
QY 198 EKVR-TGHNSSPLYKRDEESGAALSNVDAVOQLOKTPASKILLGMPTVGSRFTLAS 256  
  
D271 lhdlrappaagsnvva-vddgsml-tymrliriyivqrattynatlgdygcysgmisl 328  
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
QY 257 DTRVAPARTGSCTPGPFPRKEGGMLATIVYCVSKAKIKOR-IQQDV-VP-YIFRDNMVGF 313  
  
D329 ddtqsvrnkvnyvkrgyllgyffaw 352  
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
QY 314 DDVESFRKTYSYLKGKLGGAMVV 337
```

```
RESULT 13  
ENTRY S47133 #type complete  
TITLE chitinase (EC 3.2.1.14) - fungus (Trichoderma harzianum)  
ORGANISM 23-Nov-1994 #sequence_revision 03-Aug-1995 #text_change  
DATE 10-May-1996
```

```
ACCESSIONS  
S47133  
REFERENCE  
#authors Carosello, C.; Gutierrez, A.; Jimenez, B.; van Montagu, M.;  
Herrera-Strella, A.  
#submission submitted to the EMBL Data Library, June 1994  
#description Characterization of ech42, a Trichoderma harzianum  
endochitinase gene expressed during mycoparasitism.  
#accession S47133 preliminary  
#status Preliminary  
#molecule_type DNA  
#residues 1-424 #label CAR  
#cross_references EMBL:X79381
```

```
GENETICS  
#introns 48/2; 81/2; 98/1  
CLASSIFICATION #superfamily Streptomycetes chitinase chl40  
KEYWORDS glycosidase; hydrolase  
SUMMARY #length 424 #molecular_weight 46298 #checksum 8773
```

```
Query Match 21.4%; Score 582; DB 5; Length 424;  
Best Local Similarity 36.0%; Pred. No. 4,43e+80;  
Matches 102; Conservative 69; Mismatches 95; Indels 17; Gaps 13;
```

```
I12 glfk-lkkanrnlkmisigvwst-nfpasastdanrkfahtailmkdwfgldlv 169  
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
```

QY	57	QEFNLKRNKMKLTLLAIGWNGCTQKFDMDVATANNROTFPVNSAIRLRKRSFQCLD	116
Db	170	dweyp-addtqatmvl1lkeirsqldayaayg-agp-yhflislaapegpehy-sfilhm	225
QY	117	DWEYSGSGSPAVDKERFTTLVQDLANAFQEOEAQTSKERLLLSAAVPAQTYVDAGYEV	176
Db	226	sdlqgvlidyvnlmaydyagwsygsyghdanlfampsnpsp-yntcgaikdy1kqgvpa	284
QY	177	DKIAQNLDFVNLMAVDHFGSWEKTVGHNSPLYKRQESGAASLNDVAAVOQWLQKTPA	236
Db	285	skivlmpilygraf-estgg-1qgtysgigs-gsw-englwdy-kvlpkagatcyd	336
QY	237	SKLIGMTYGRSTTLASSSDTRVGAPATGSCITGPPTKEGMLATYEVCSWKGATKQ--	294
Db	337	svagayysydpsskelisfdtrpdmntekvys1knl1q1gsgmfw	379
QY	295	RIQDQKVPIFRDNQWGFDDVESFKTKVSYLKQKIGLGAMVW	337
RESULT	14		
ENTRY		QJ1975	#type complete
TITLE		chitinase (EC 3.2.1.14) 1 precursor - imperfect fungus	
ORGANISM		(aphanocladium album)	
DATE		03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change	
ACCESSIONS		QJ1975	
REFERENCE		QJ1975	
authors		Blaiseau, P.L.; Lafay, J.F.	
#journal		Gene (1992) 120:243-248	
#title		Primary structure of a chitinase-encoding gene (chl1) from the filamentous fungus Aphanocladium album: similarity to bacterial chitinases.	
#accession		QJ1975	
#molecule_type		mRNA	
##residues		1-423 ##label B1A	
##cross-references		GB:X64104	
COMMENT		This enzyme is essential for the degradation of insoluble chitin.	
GENETICS			
#gene		chl1	
#introns		47/2; 80/2; 97/1	
CLASSIFICATION		#superfamily Streptomyces chitinase chl40	
KEYWORDS		glycosidase; hydrolase	
FEATURE			
1-34		#domain signal sequence #status predicted #label sig\	
35-423		#product chitinase 1 #status predicted #label MAT	
SUMMARY		#length 423 #molecular_weight 46072 #checksum 8802	
Query Match		21.2%; Score 576; DB 5; Length 423;	
Best Local Similarity		36.2%; Pred. No. 5.50e-79;	
Matches 101; Conservative		65; Mismatches 95; Indels 18; Gaps 16;	
Db	115	lkkgrnmkvnlsiggywst-nrpaassatctktagsavgfkmkgfdgdldideyp	173
QY	62	LKKMPKTKTLTLAGGWNFGQKFTDMVATANNROTFPVNSAIRLRKRSFQCLDDEYR	121
Db	174	-adatqagmvl1lqavrselidsyaayg-akg-rhflislaapeghnyk1kfa-e-1g	228
QY	122	GSQGSPPAVDKERFTTLVQDLANAFQEOEAQTSKERLLLSAAVPAQTYVDAGYEV	180
Db	229	kvldylnlmaydyagswnytcghdanlyanpnpnatp-yntcdavgay1ngvpankiv	287
QY	181	QNLDFVNLMAVDHFGSWEKTVGHNSPLYKRQESGAASLNDVAAVOQWLQKTPA	240
Db	288	lmpilygrsfq-qc-e-gg-1qgtyngigs-gsw-englwdykalp-agaatyvcdttak	339
QY	241	IGMPYGRSFTLASSSDTRVGAPATGSCITGPPTKEGMLATYEVCSWKATKRIODOK	300
Db	340	gcysydpstkelisfdtrpdmntekvys1kng1q1gsgmfw	378
QY	301	VPIR-RD-NQWGFDDVESFKTKVSYLKQKIGLGAMVW	337









```

Db      83 lknmpulkltsvgyumfsgqrtskkslasnsgqrtskkslvgpvlrlfghgdldlmlp 142
      ||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Qy      62 LKKMNPKLITLALGGNNFGQKTTDMWALNNQKTVNSIRFLRKXYSFDGDLDMEXR 121
      ||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      143 grt-----dkqhtftllkemmkaefikeaq-pgkqjlllsaalasgkvrltssydlahtsq 196
      ||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Qy      122 GSGGSPAVMDKRFPTTIVODLANARQGEAQISGKRERLLLSAAYPRAGQRYVDAGYEVDKIAQ 181
      ||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      197 hldfslmtydflnqawrgtgcghnspflfrggedaspdfsntdyavgymlllgaspakslvm 256
      ||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Qy      182 NLDEVNLMAYDYFHSGSWEKVTGHNSPILKPKRGEESGAASLNVDAVQWMLQKTPASRLIL 241
      ||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      257 giprfgrfstllass-etsvgagpisgpglprfctkeagtlaiyeicdlirgatvhrtlrgq 315
      ||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Qy      242 GMPYRGREFTLTIASSDPRVGAAPAGSGSPGPFTTEGGEMALYYEVCSSV-KGATKORIDOK 300
      ||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      316 vpyatkgawgwgyddgsesvskxykylldrqldqagmawmalldldqgsfgcgqdlrfpftna 375
      ||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Qy      301 VPIYFRDQWQWGFDDVESFRTKVSYLKOKGGLGAMVWALLDLDPAGSFCNQG-RYPLIQT 359
      ||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      376 lkdaa 381
      ||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Qy      360 LRQELS 365
      ||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

```

ID	RESULT	2	STANDARD;	PRT;	504 AA.
AC	CHIT1	BRUMA			
AC	P29030;				
DT	01-DEC-1992	(REL. 24, CREATED)			
DT	01-DEC-1992	(REL. 24, LAST SEQUENCE UPDATE)			
DT	01-FEB-1995	(REL. 31, LAST ANNOTATION UPDATE)			
DE	ENDOCHITININASE PRECURSOR	(EC 3.2.1.14) (MFI ANTIGEN).			
OS	BRUGIA MALAYI.				
OC	EUCARYOTA; METAZOA; ACCELOMATES; NEMATODA; SECERNENTEA; SPIRURIDA.				
RN	[1]				
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.				
RX	MEDLINE; 92179220.				
RA	FURUKAWA J.A., LANE W.S., SMITH R.F., PLESSENS W.F., PERLER F.B.,				
RL	FOURMAN J.A., ACAD. SCI. U.S.A. 89:1548-1552(1992).				
CC	-1- FUNCTION: THE MFI ANTIGEN IS A MICROFILARIAL CHITININASE, WHICH MAY				
CC	FUNCTION TO DEGRADE CHITIN-CONTAINING STRUCTURES IN THE MICRO-				
CC	FILARIA OR IN ITS MOSQUITO VECTOR DURING PARASITE DEVELOPMENT AND				
CC	TRANSMISSION.				
CC	-1- CATALYTIC ACTIVITY: HYDROLYSTS OF THE BETA-1,4 LINKAGES OF				
CC	N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.				
CC	-1- PTM: O-GLYCOSYLATED.				
CC	-1- KNOWN TO BIND CALCIUM.				
CC	-1- DEVELOPMENTAL STAGE: THE APPEARANCE OF THE MFI ANTIGEN CORRESPOND				
CC	WITH THE ONSET OF THE PARASITE'S ABILITY TO INFECT THE MOSQUITO.				
CC	-1- SIMILARITY: BELONGS TO CHITININASE CLASS II (FAMILY 18 OF GLYCOSYL				
CC	HYDROLASES).				
DR	EMBL; M73689; G156064; -.				
DR	PIR; A38221; A38221.				
DR	PROSITE; PS01095; CHITININASE_18.				
KW	HYDROLASE; GLYCOSIDASE; CHITIN DEGRADATION; SIGNAL; ANTIGEN;				
KW	REPEAT; GLYCOPROTEIN; CALCIUM-BINDING.				
FT	SIGNAL	1	22		
FT	CHAIN	23	504	ENDOCHITININASE.	
FT	DOMAIN	23	400	CATALYTIC.	
FT	DOMAIN	401	450	SER/THR-RICH (LINKER).	
FT	DOMAIN	407	448	3 X 14 AA APPROXIMATE TANDEM REPEATS.	
FT	ACT_SITE	148	148	PROTON DONOR (BY SIMILARITY).	
SQ	SEQUENCE	504 AA;	55971 MM;	4DAY5EB CRC32;	
Best Match	Onsery March	39.9%;	Score 1086;	DB 2;	Length 504;
Matches	158; Conservative	42.8%;	Pred. No. 6.2e-219;		
		94;	Mismatches 102;	Indels	15; Gaps
					11.

D	b	87	avcklctngjklkllslyggynjsafalijaksagctrefifsafalirknmffordld	146
O	y	58	EFNCLKMMNPKLTLAIGGNNGTQKFTLVAVANRRQFVSALRELFKYSFDDLDL	117
D	b	147	wepv-v-gv-a-e-eh-aklveamktafveaaktsqqrlllcaavsaqgtldgsyve	200
O	y	118	WEYSGSGSPAVDERRTTLVODLANNFOQEAOTSGERLLLSAAYPAGQTYDAGYEVD	177
D	b	201	slgknfllfimsydlhgsweknhvdlhghklhptkgevsigifnfeaaqywasckmpke	260
O	y	178	KIAONLDFVNLMADFHGSEWKEVTGHSNPLYRKQEESSGAASLNVDAAYQOMLOKGPAS	237
D	b	261	klllgfpmayagvwlndpsetalgaasrpsasaklnpagtasyweickylkegketyv	320
O	y	238	KLIIGMPTYGSEFTLASSSDTRVAGAPATGSGTGPFPKKEGGMALAYEVC5W-KGATKORI	286
D	b	321	hgqevaymwkgddvwydydneetrlmmkvlhektygsafinaldfadfigksckgkypv	380
O	y	297	QDQKV-PYFRDNQWGFDDVESEKTVSYLKOKGLGAGVAMVALLDDDFAGFCNCGRYP	355
D	b	381	llnaisseel 389	
O	y	356	LIQTLRQEL 364	

ID	RESULT_3		STANDARD;	PRT; 554 AA.
AC	CHIT_MANE			
DJ	p3636z,			
DT	01-JUN-1994 (REL. 29, CREATED)			
DT	01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)			
DE	01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)			
OS	ENDOCHITTINASE PRECURSOR (EC 3.2.1.14).			
OC	MANDUCA SEXTA (TOBACCO HAWKMOOTH) (TOBACCO HOMOMORF), EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; LEPIDOPTERA. [1]			
RN	[1] SEQUENCE FROM N.A.			
RP	MEDLINE; 93357793.			
RA	KRAMER K.J., CORPUZ L., CHOI H.K., MUTHUKRISHNAM S.; INSECT BIOCHEM. MOL. BIOL. 23:691-701(1993).			
CC	- FUNCTION: DIGEST CHITIN IN THE EXOSKELETON DURING THE MOULTING PROCESS.			
CC	- CATALYTIC ACTIVITY: HYDROLYSIS OF THE BETA-1,4 LINKAGES OF N-ACETYL-D-GALACTOSAMINE POLYMERS OF CHITIN.			
CC	- DEVELOPMENTAL STAGE: HIGH LEVELS SEEN IN THE EPIDERMIS ON DAY 0, BUT RAPIDLY DISAPPEARS AND IS UNDETECTED ON DAYS 1-4 OF FIFTH INSTAR. IT REAPPEARS ON DAY 5 AND PEAKS ON DAY 7 AFTER WHICH A RAPID DECLINE IS SEEN. IN THE GUT IT IS DETECTED ON DAY 6 WITH LOWER LEVELS SEEN ON DAYS 0, 7 AND 8.			
CC	- TISSUE SPECIFICITY: EPIDERMIS AND GUT.			
CC	- SUBCELLULAR LOCATION: SECRETED.			
CC	- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL HYDROLASES).			
DR	EMBL; U02270; g406049; -; EMB1; s64757; -, NOT_ANNOTATED.CDS.			
DR	HSSP; P07254; ICTN.			
DR	PROSITE; PS01093; CHITINASE_18.			
KW	HYDROLASE; GLYCIDASE; CHITIN DEGRADATION; SIGNAL; GLYCOPROTEIN.			
FT	SIGNAL 1 19 POTENTIAL.			
FT	CHAIN 20 554 ENDOCHITTINASE.			
FT	DOMAIN 396 453 SER/THR-RICH.			
FT	ACT_SITE 146 146 PROTON DONOR (BY SIMILARITY).			
FT	CARBONYD 85 85 POTENTIAL.			
FT	CARBONHD 303 303 POTENTIAL.			
FT	CARBOND 407 407 POTENTIAL.			
FT	CARBONYD 545 545 POTENTIAL.			
SO	SEQUENCE 554 AA; 62203 MW; FA8/F8AD CRC32;			
Bd	Query Match 37.3%, Score 1015; DB 2; Length 554;			
Bd	Best Local Similarity 40.4%; Pred. No. 4,86e-202;			
Matches	155; Conservative 98; Mismatches 107; Indels 24; Gaps 18;			
Dj	atcygfsnvaayrpgvgrgyiedipvektchittysfigtgensevlldpdeldvdngk 82  ::   :      :  :     ::  :  :  :  :			

Qy	1	AKVCFEFNNQAVYQNGSARFLPKRLDSDLSCTHLILYANAGNTN-H-Q-LSR-TRMN-DETL	55
Dd	83	frnfstlsrshpsvkfmvavgwgvaegsskxshmvagqstlmsfirsvsflkxyddgdld	142
Qy	56	YOEENGKLKPKMKLTLLATLALNMGFCQKFTMDVATANNQOTFVNSAIRLFRKPSDGD	115
Dd	143	ldwvppgaadrggfsdkdkfilyvgeltraf---lrv-gkygel-taavplanfimg	197
Qy	116	LDMEYPSG--GGSAPVAKERETTLIVQDLANAFQOEACTSKEKHLILSAAPAGQIVDAG	173
Dd	198	yhvpelcgealdhvmaydyldrgnavgafadvshpkyrphqvaeyeklnvndghlwekyg	257
Qy	174	YEVDKINQNDLFPVLMAYDPHGSMBEXKTYGHNSPLYKKGDESGAASLNDPAVQQLONG	233
Dd	238	cpshkivvgjrlfygrsftlsagunnnylgfinkkaaggqdpabytnatgfwayelctev	317
Qy	234	TPASKLITGMPTLYGRSFLLASSSDT-RVGA--P-ATSGSGPGPFTEKGGLAYEVCSS--	287
Dd	318	dkdsgwtkkkdeegskcpoyekygtqvwgyedpprsvelkmwllkqkylgamtalmdfd	377
Qy	288	WKG-A--TKQRIQOQKVPYFRDQWGWGDFVESFKTYKYLKQKGLGAMVWALDLDF	344
Dd	378	qgl--cge-knp1lk1klnhks--sy 398	
Qy	345	AGFSGNGRYPLTQTLRQELSLRP 368	
RESULT	4	STANDARD; PRT; 699 AA.	
ID	CH11_BACCI		
AC	P20533;		
DT	01-FEB-1991 (REL. 17, CREATED)		
DT	01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)		
DT	01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)		
DE	CHITINASE A1 PRECURSOR (EC 3.2.1.14).		
GN	CHAI.		
OS	BACILLUS CIRCULANS.		
OC	PROKARYOTA; FIRMITICES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE		
RP	[1]		
RN	SEQUENCE FROM N.A.		
RC	STRAIN-WL-12.		
RX	MEDLINE: 90368776.		
RA	WARANABE T., SUZUKI K., OYANGI W., OHNISHI K., TANAKA H.;		
RL	J. BIOL. CHEM. 265:15659-15665(1990).		
RN	[2]		
RP	MOTAGENESIS.		
RC	STRAIN-WL-12.		
RX	MEDLINE: 93366760.		
RA	WARANABE T., KOHORI K., MIYASHITA K., FUJII T., SAKAI H.,		
RA	UCHIDA M., TANAKA H.;		
RL	J. BIOL. CHEM. 268:18567-18572(1993).		
CC	-1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE BETA-1,4 LINKAGES OF		
CC	-1- N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.		
CC	-1- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS		
CC	-1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL		
CC	HYDROLASES).		
DR	EMBL: M57601; G142688; .		
DR	PIR: A38368; A38368.		
DR	HSSP: P07254; ICTN.		
DR	PROSITE: PS01095; CHITINASE_18.		
KW	HYDROLASE; GLYCOSIDASE; CHITIN DEGRADATION; SIGNAL; REPEAT.		
FT	SIGNAL	1	41
FT	CHAIN	42	699
FT	DOMAIN	42	460
FT	DOMAIN	465	549
FT	DOMAIN	560	644
FT	ACT_SITE	204	204
FT	MUTAGEN	200	200
FT	MUTAGEN	200	200
FT	MUTAGEN	204	204
SQ	SEQUENCE	699 AA;	73677 MW; 5A1777CC CRC32;
Query Match	25.7%;	Score 698;	DB 2; Length 699;
Best Local Similarity	40.1%;	Pred. No. 1.82e-127;	

Matches	119;	Conservative	73;	Mismatches	85;	Indels	20;	Gaps	18;
Db	144	qlnk kqctfnpl kllisvgwtw--snifsdvaataatrevfanaavdflrkynfdgvd	202						
Qy	58	EFNGLKKNPKTKTLTAIGWNGFGQKQFTDMVATANNRQTFVNSAIRLKRYSFSDGLDLD	117						
Db	203	weyprsggladgnskrpdknytl llsklrk laaavdvk-kyl ltiaagaatya-a	260						
Qy	118	WEYP--GS-QG-S--PADVKEFFTLVODLANAFQOEAOISCKEKLTL SAAPACOTYVDA	172						
Db	261	ntelakiaivdw inimfydngawqk sahnap lunydpaaasagvydantfvaagag	320						
Qy	173	GTEVDKIQONIDFYNLAAYDFHGSNEKTYGNISPL-YKRPD-ESGAA-A-SLNDAAVQO	228						
Db	321	hldagvpaakiv gvpyfyrgwd-gcaagag--ggydctcgssvsgtveasfdfyd ean	378						
Qy	229	W QKQTPASKLILGMPTYGRSFTLASSSDTRVGA PAGGSGRPGPFTREGGLAYEV--CS	287						
Db	379	ylnkgyrryundekavpyl naenkr tsydsdaesvyk kaylksq ggamfvel	435						
Qy	288	W--K-GATKORIODQKVPYIFRD-NQW-VGDDVSESFRTKSYLKOKGLGAGAWMAL	339						
RESULT 5 STANDARD: PRT: 423 AA.									
ID	CH11	APHAL							
AC	P32470.								
DT	01-OCT-1993	(REL. 27, CREATED)							
DT	01-FEB-1996	(REL. 33, LAST SEQUENCE UPDATE)							
DT	01-FEB-1996	(REL. 33, LAST ANNOTATION UPDATE)							
DE	CHITINASE 1	PRECURSOR (EC 3.2.1.14).							
GN	CH11.								
OS	APANOCCLADIUM ALBUM.								
OC	EUTAROTA; FUNGI; DEUTEROMYCOTINA (IMPERFECT FUNGI).								
NC	[1]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN-ETHM 483.								
FX	MEDLINE; 93013040.								
RA	BLAISEAU P.-L., LAEAY J.-F.;								
RL	GENE 120:243-248(1992).								
RN	[2]								
RP	SEQUENCE OF 35-57.								
FX	MEDLINE; 92164637.								
RA	BLAISEAU P.-L., KUNZ C., GRISON R., BERTHEAU Y., BRYGOO Y.;								
RL	CURR. GENET. 21:61-66(1992).								
CC	-I- CATALYTIC ACTIVITY: HYDROLYSIS OF THE BETA-1,4 LINKAGES OF								
CC	N-ACETYL-D-GUCOSAMINE POLYMERS OF CHITIN.								
CC	-I- SUBCELLULAR LOCATION: SECRETED								
CC	-I- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL								
CC	HYDROLASES).								
DR	EMBL; X64104; G429026; -.								
DR	PIR; J01975; J01975.								
DR	HSSP; P07254; ICDN.								
DR	PROSITE; PS01095; CHITINASE_18.								
KW	HYDROLASE; GLYCOSIDASE; CHITIN DEGRADATION; SIGNAL; ZYMOGEN.								
FT	SIGNAL	1	22	POTENTIAL.					
FT	PROPEP	23	34	POTENTIAL.					
FT	CHAIN	35	423	CHITINASE 1.					
FT	ACG SITE	171	171	PROTON DONOR (BY SIMILARITY).					
SO	SEQUENCE	423 AA;	46058 MW;	88456E6 CRC32;					
Query Match 21.2%; Score 576; DB 2; Length 423;									
Best Local Similarity 36.2%; Pred. No. 2,50e-99;									
Matches 101; Conservative 65; Mismatches 95; Indels 18; Gaps 16;									
Db	115	lkkqrrmmkwnl sgywtwst-n paaasaaat k fagaavg fmkdw gfgid ideyp	173						
Qy	62	LKKMPKTKTLTAIGWNGFGQKQFTDMVATANNRQTFVNSAIRLKRYSFSDGLDLD	121						
Db	174	-adatqagmw l qavse dsyaayg-akg-nh fl slaapgd nyn k fa-e- lg	228						
Qy	122	GSQGS P AVDKER FTTLVODLANA QOE QTS GKER LLISA AVP AG-OT YD AG EVD KIA	180						
Db	229	kvl d in may dyg aw sn y c gh an y an qp n atp-n v td av gay ng v pan k iv	287						

181 QNDFVNLMAYPDEHSGMEKVTGHNPSLYKROESGAALSNVDAVQOMLOKGPASKL 240  
Db 288 lmpdygrsfq-qt--eg-igkpyngfqs-gsw--engivdykalpk-asatkcddtak 339  
QY 241 LGMPTGRSTTLASSSDTRGAPATGSGTGPTEKGMALAYEVCWSMGATKORIODOK 300  
Db 340 gcysydpstkelstfclpamistekvsklgkqigntmf 378  
QY 301 VPYIF-RD-NQWGFDDVESFKTKVSYLKOKGLGAMVW 337

RESULT 6  
ID CHIA-TRIHA STANDARD; PRT; 423 AA.  
AC P48827;  
DT 01-FEB-1996 (REL. 33, CREATED)  
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
DE 42 KD ENDOCHITINASE PRECURSOR (EC 3.2.1.14).  
GN CHIT42.  
OS TRICHODERMA HARZIANUM.  
OC EUKARYOTA; FUNGI; DEUTEROMYCOTINA (IMPERFECT FUNGI).  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 35-52; 93-107; 371-385 & 397-414.  
RX MEDLINE; 95269313.  
RA GARCIA I., LORA J.M., LA CRUZ J., BENITEZ T., LOBELLE A.,  
RA PINTOR-TORO J.A.;  
RL CURR. GENET. 27:83-89(1994).  
CC -1- FUNCTION: MORPHOGENETIC ROLE DURING APICAL GROWTH/CELL DIVISION  
AND DIFFERENTIATION (CELL WALL MORPHOGENESIS). ANTIFUNGAL  
AGENT.  
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE BETA-1,4 LINKAGES OF  
N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.  
CC -1- SUBCELLULAR LOCATION: SECRETED.  
CC -1- INDUCTION: SPECIFICALLY INDUCED BY CHITIN AND IS CATABOLITE  
REPRESSOR.  
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
HYDROLASES).  
CC EMBL; 578423; G999376; -.  
DR HYDROLASE; GLYCOSIDASE; CHITIN DEGRADATION; SIGNAL; ZYMOGEN;  
KW CHITIN-BINDING.  
FT SIGNAL 1 22 POTENTIAL.  
FT PROPEP 23 34  
FT CHAIN 35 423 42 KD ENDOCHITINASE.  
FT ACT SITE 171 171 PROTON DONOR (BY SIMILARITY).  
FT CARBOHYD 218 218 POTENTIAL.  
SQ SEQUENCE 423 AA; 46056 MW; 899DA50A CRC32;

Query Match 20.6%; Score 561; DB 2; Length 423;  
Best Local Similarity 35.7%; Pred. No. 6.68e-96;

Matches 101; Conservative 71; Mismatches 94; Indels 17; Gaps 13;

Db 111 qltk-vtkarqgkvlvlsigvwt-nfpaastadnknfaktatctfkdvgfgdtd 168  
QY 57 QENNGKAKMNPKLTLLAIGMNFQKFTDMVATANNROTFTVNSAIRFKRISFDL 116  
Db 169 dweyp-adatqasnmlllkevsqrdaaay-apg-yhflitlaapagkdy-sklr 224  
QY 117 DWYPSGSPAYDKERTTLVODLANAFQOEQOTSKEKRLLSAAYPAQOTVVDAGEY 176  
Db 225 adigvldylnlnaydagsfsltdhnanlfnpsnpatp-fntdsavkdyingvpa 283  
QY 177 DKIAQNLDFVNLMAYPDEHSGMEKVTGHNPSLYKROESGAALSNVDAVQOMLOK 236  
Db 284 nkivlmplygrsf--qntag--igqlyngvgs-gsw--eagivdykalpk-asatvqd 335  
QY 237 SKLILGMPYGRSTTLASSSDTRGAPATGSGTGPTEKGMALAYEVCWSMGATKORI 296  
Db 336 svakgyysynsakelstfclpamistekvsklgkqigntmf 378  
QY 297 ODOKVPIFRD-NQWGFDDVESFKTKVSYLKOKGLGAMVW 337

RESULT 7  
ID CHIL COCIM STANDARD; PRT; 427 AA.  
AC P54196;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE ENDOCHITINASE 1 PRECURSOR (EC 3.2.1.14) (COMPLEMENT-FIXATION ANTIGEN)  
DE (CF-ANTIGEN) (CF-AG).  
GN CT51.  
OS COCCIDIOIDES IMMITIS.  
OC EUKARYOTA; FUNGI; DEUTEROMYCOTINA (IMPERFECT FUNGI).  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN-C735;  
RC MEDLINE; 96144270.  
RA PISHKO E.J., KIRKLAND T.N., COLE G.T.;  
RL GENE 167:173-177(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SILVEIRA;  
RA YANG C., ZHU Y., MAGEE D.M., COX R.A.;  
RL SUBMITTED (MAR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.

CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE BETA-1,4 LINKAGES OF  
N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.  
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
HYDROLASES).  
DR EMBL; 141663; G1200190; -.  
DR EMBL; U51271; G1256769; -.  
DR EMBL; U33265; G1255728; -.  
KW HYDROLASE; GLYCOSIDASE; CHITIN DEGRADATION; CHITIN-BINDING; SIGNAL;  
KW GLYCOPROTEIN.  
FT SIGNAL 1 2 ? POTENTIAL.  
FT CHAIN 387 427 ENDOCHITINASE 1.  
FT CARBOHYD 387 387 POTENTIAL.  
FT CONFLICT 15 47 RWLSRLCLELGEKREMTLSTVATVTTDIO -> VQAS  
FT CONFLICT 159 199 SMSMNPYVPEAGEGFSRVYFVW (IN REF. 2).  
SQ SEQUENCE 427 AA; 47629 MW; 6CB9AF73 CRC32;

Query Match 20.1%; Score 546; DB 2; Length 427;  
Best Local Similarity 34.1%; Pred. No. 1.75e-92;

Matches 95; Conservative 71; Mismatches 94; Indels 19; Gaps 17;

Db 115 lkknnnllkllsfgvty-spnfktpasteegrkfadtstlkmldyfgddldwep 173  
QY 62 LKMANKLTLLAIGMNFQKFTDMVATANNROTFTVNSAIRFKRISFDL 121  
Db 174 ede-kqandfvlllkacrealdaysak-hpngk-kflltlaapagpny-nk-1klaend 228  
QY 122 GSGSPAYDKERTTLVODLANAFQOEQOTSKEKRLLSAAYPAG-QTVVDAGEYDKIA 180  
Db 229 kyldfnlnaydagsvsgvmsnfvsttkp-estfcsdkavkdyikagvpankiv 287  
QY 181 QNDFVNLMAYPDEHSGMEKVTGHNPSLYKROESGAALSNVDAVQOMLOKGPASKL 240  
Db 288 lmpdygrsf--ast-dg--lqtsfngvg-gsw--engivdykalpk-asatvleledia 339  
QY 241 LGMPTGRSTTLASSSDTRGAPATGSGTGPTEKGMALAYEVCWSMGATKORIODOK 300  
Db 340 asysydknkryllsydgtvklagkkaeyilkngmggm-w 377  
QY 301 VPYIF-RDQW-VGFPDDVESFKTKVSYLKOKGLGAMVW 337  
RESULT 8  
ID CHIA-SERMA STANDARD; PRT; 563 AA.  
AC P07254;  
DT 01-APR-1988 (REL. 07, CREATED)  
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)  
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
DE CHITINASE A PRECURSOR (EC 3.2.1.14).  
GN CHIA.  
OS SERATIA MARCESCENS.

[illegible]



Db 425 gsdgsk-idaa-dygeaskyidkynmtygfifgawak-ngtaphnplitaydglpqgfn 481  
QY 163 VPGQTYVDGYEVNDKIAQNLDFVNLMAYPDHSWENVTGSHNSPILKRRQESGAAA-SLN 221  
Db 482 tadamakfkxagvpadklllgfygrgvt-gvtqsap-gtatagpat-gty--eaagied 536  
QY 222 VDAAYQOMLQKTPASKLLIGMPTYGSRFTLASSSDTRVGAAPATGSGTTPGPTKEGGM 281  
Db 537 y-kvlnscpatgtlag--tayahcgsmwsvydtipaliksmgwaegqglggaftweifsg 593  
QY 282 YVEVCSMKCATKQRIQDQKVPYIFRDNMQWGFDDVESFKTKVSYLKQKGLGAMVWALDL 341  
Db 594 dtang 598  
QY 342 DDFAG 346

RESULT 13  
ID CHIT-STRPL STANDARD; PRT; 610 AA.  
AC P11220;  
DT 01-JUL-1989 (REL. 11, CREATED)  
DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)  
DE 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
DE CHITINASE 63 PRECURSOR (EC 3.2.1.14).  
GN CHTA  
OS STREPTOMYCES PLECATUS.  
OC PROMARYOTA; FIRMICUTES; ACTINOMYCETALES; STREPTOMYCETACEAE.  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 92192480.  
RA ROBBINS P.W., OVERBYE K., ALBRIGHT C., BENFIELD B., PERO J.;  
RL GENE 111:69-76(1992).  
RN (2)  
RP SEQUENCE OF 1-45 FROM N.A., AND SEQUENCE OF 31-45.  
RX MEDLINE: 88087127.  
RA ROBBINS P.W., ALBRIGHT C., BENFIELD B.;  
RL J. BIOL. CHEM. 263:443-447(1988).  
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE BETA-1,4 LINKAGES OF  
CC -1- N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.  
CC -1- INDUCTION: BY CHITIN.  
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
CC HYDROLASES).  
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE IIT-LIKE DOMAIN.  
CC -1- SIMILARITY: CONTAINS A BACTERIAL-TYPE CELLULOSE-BINDING DOMAIN  
CC (CBD).  
DR EMBL: M82804; G153216; -;  
DR EMBL: M18397; G153209; -;  
DR PIR: A29912; A29912;  
DR PIR: JH0573; JH0573;  
DR HSSP: P07234; ICTN.  
DR PROSITE: PS00018; EF\_HAND.  
DR PROSITE: PS00561; CBD\_BACTERIAL.  
DR PROSITE: PS01095; CHITINASE\_18.  
KW HYDROLASE; GLYCOSIDASE; CHITIN DEGRADATION; CHITIN-BINDING; SIGNAL.  
FT SIGNAL 1 30  
FT CHAIN 31 610 CHITINASE 63.  
FT DOMAIN 35 140 CELLULOSE-BINDING.  
FT DOMAIN 148 229 FIBRONECTIN TYPE-III.  
FT DOMAIN 236 610 CATALYTIC.  
FT ACT\_SITE 383 383 PROTON DONOR (BY SIMILARITY).  
FT CONFLICT 3 3 F -> I (IN REF. 2).  
SQ SEQUENCE 610 AA; 63974 MW; 2F5E8E35 CRC32;

Query Match 10.5%; Score 287; DB 2; Length 610;  
Best Local Similarity 26.4%; Pred. No. 1,72e-35;  
Matches 78; Conservative 63; Mismatches 134; Indels 21; Gaps 16;

Db 320 fnglrmkkaeyphklllysfgwtsvsg-fpdavknpaafakschdlvedprwadvfdgl 378  
QY 56 YQEFNGLKKNPKIKTLAIGWNFGQKFTDMVATFANNQTFVNSAIRFLRKYs-FDGL 114  
Db 379 dlwepnaeglsdelsapnatsmmkamraefgqdyllta-a---vtadsgdgsk-id 433

QY 115 DLDMEYPSQGS-S--PAVDKERTTIVODLANFQCEAOTSNGERILLISAAPAGCTIYD 171  
Db 434 aa-dygeaskyidkynmtydfffagawak-ngtaphspinaydgllpqgfttadamakfk 491  
QY 172 AGEVVDKIAQNLDFVNLMAYPDHSWENVTGSHNSPLYKRRQESGAAA-SLNDAAVQOML 230  
Db 492 skvpadklllgfygrgvt-gvtqsap-gtatagpa-gty--eaagledy-kvlnsc 545  
QY 231 QKTPASKLLIGMPTYGSRFTLASSSDTRVGAAPATGSGTTPGPTKEGGMAYEVC5MWK 290  
Db 546 patgvaag--tayahcgsmwsvydtipaliksmgwaegqglggaftweifsgdtng 599  
QY 291 ATKQRIQDQKVPYIFRDNMQWGFDDVESFKTKVSYLKQKGLGAMVWALDDDFAG 346

RESULT 14  
ID KTXA\_KLULA STANDARD; PRT; 1146 AA.  
AC P09805;  
DT 01-MAR-1989 (REL. 10, CREATED)  
DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)  
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
DE KILLER TOXIN ALPHA AND BETA SUBUNITS PRECURSOR (RF2 PROTEIN)  
DE (ENOCHITINASE (EC 3.2.1.14)).  
OS KLUYEROMYCES LACTIS (YEAST).  
OC PLASMID PGK1-1.  
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN-CBS 2359/152;  
RA SOR F., FUKUHARA H.;  
RL CORR. GENET. 9:147-155(1985).  
RN (2)  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 84297209.  
RA STARK M.J.R., MILEHAM A.J., ROMANOS M.A., BOYD A.;  
RL NUCLEIC ACIDS RES. 12:6011-6030(1984).  
RN (3)  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 85037931.  
RA TOKUNAGA M., KAWAMURA A., HISHINUMA F.;  
RL NUCLEIC ACIDS RES. 12:7581-7597(1984).  
RN (4)  
RP IDENTIFICATION OF PROTEIN, AND SEQUENCE OF 30-44 AND 895-916.  
RX MEDLINE: 87004569.  
RA STARK M.J.R., BOYD A.;  
RL EMBO J. 5:1995-2002(1986).  
RN (5)  
RP SIMILARITY TO CHITINASE OF ALPHA-SUBUNIT.  
RX MEDLINE: 90259069.  
RA BRADSHAW H.D. JR.;  
RL NATURE 345:299-299(1990).  
RN (6)  
RP CHITINASE ACTIVITY OF ALPHA-SUBUNIT.  
RX MEDLINE: 91301161.  
RA BUTLER A.R., O'DONNELL R.W., MARTIN V.J., GOODAY G.W., STARK M.J.R.;  
RL EUR. J. BIOCHEM. 199:483-488(1991).  
CC -1- FUNCTION: THE ALPHA SUBUNIT IS A POTENT EXOCHITINASE. ALONG WITH  
CC THE BETA SUBUNIT IT PLAYS A ROLE IN THE INITIAL INTERACTION OF THE  
CC TOXIN WITH SENSITIVE CELLS AND ALLOW THE GAMMA SUBUNIT (THE ACTIVE  
CC TOXIN) TO GAIN ENTRY INTO THE CELL.  
CC -1- PTM: RF2 IS POTENTIALLY SPLIT BY MEMBRANE-BOUND BASIC AMINO ACID-  
CC SPECIFIC PEPTIDASE TO YIELD THE ALPHA AND BETA SUBUNITS.  
CC -1- SUBUNIT: THE KILLER TOXIN IS COMPOSED OF THREE SUBUNITS: ALPHA,  
CC BETA AND GAMMA.  
CC -1- SIMILARITY: LOCAL, TO OTHER CHITIN-BINDING PROTEINS.  
CC -1- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.  
DR EMBL: X07127; G28429; -;  
DR EMBL: X00762; G28449; -;  
DR EMBL: X01095; G28499; -;  
DR PIR: S07915; S07915;  
DR HSSP: P02877; 1HEV.  
DR PROSITE: PS00026; CHITIN-BINDING.

DR PROSITE: PS01095; CHITINASE\_18.  
 KW PLASMIN; TOXIN; GLYCOPROTEIN; SIGNAL; HYDROLASE; GLYCOSIDASE;  
 KM CHITIN DEGRADATION.  
 FT SIGNAL 1 17 POTENTIAL.  
 FT PROPEP 18 29 POTENTIAL.  
 FT CHAIN 30 892 ALPHA SUBUNIT (PROBABLE).  
 FT CHAIN 895 1146 BETA SUBUNIT (PROBABLE).  
 FT DOMAIN 310 363 CHITIN-BINDING (BY SIMILARITY).  
 FT DOMAIN 364 635 CATALYTIC.  
 FT ACT\_SITE 495 495 PROTON DONOR (BY SIMILARITY).  
 FT CARBOHYD 771 771 POTENTIAL.  
 FT CARBOHYD 858 858 POTENTIAL.  
 FT CARBOHYD 868 868 POTENTIAL.  
 FT CARBOHYD 876 876 POTENTIAL.  
 FT CARBOHYD 995 995 POTENTIAL.  
 FT CARBOHYD 1082 1082 POTENTIAL.  
 FT CARBOHYD 1117 1117 POTENTIAL.  
 SQ SEQUENCE 1146 AA: 128936 MW: CDD1620F CRC32;

Query Match 9.7%; Score 263; DB 5; Length 1146;  
 Best Local Similarity 28.4%; Pred. No. 1.56e-30;  
 Matches 60; Conservative 58; Mismatches 74; Indels 19; Gaps 14;

Db 425 ddsafesafklyvtsek-k-lpsfggwdstpsstytftrnavkcdqnrntfannlilm 482  
 Qy 52 DETLQEFNGELKRMKRLTLAIGWNGT-Q-KFT--DWAATANNROTFTVNSAIRFL 106  
 Db 483 nkyndgldlweyepapdpdpdsssgsnlylflklkxmps-gktl-slaips 540  
 Qy 107 RKTSFGLDMEYPSGSPAVDKERTTLVODLANAFQERQTSGERLLISAAPNG 166  
 Db 541 ywylnk-fpsldqntvdyvmytydlhgw--yqkans-yl--nchtprk-eledai 592  
 Qy 167 QYVVDAGYEVNDKIAQLDFVNLIMAYDFHGSWEKVTGHSNPLYKRGESGAASLVNDAV 226  
 Db 593 km-lkagvfkfkvgvgvanygrsykmvntn 622  
 Qy 227 QQWLQK-GTPASKLILGMPTGRSFTLASS 256

RESULT 15  
 ID GP39 BOVIN STANDARD; PRT: 36 AA.  
 AC P30922;  
 DT 01-JUL-1993 (REL. 26, CREATED)  
 DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE CARTILAGE GLYCOPROTEIN-39 (GP-39) (39 KD WHEY PROTEIN) (FRAGMENT).  
 OS BOS TAURUS (BOVINE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 OC EUTHERIA; ARTIODACTYLA.  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE: 88106603.  
 RA REJMAN J.J., HURLEY W.L.;  
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 150:329-334(1988).  
 CC -I- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CAPACITY OF CELLS TO  
 CC RESPOND TO AND COPE WITH CHANGES IN THEIR ENVIRONMENT.  
 CC -I- PTM: GLYCOSYLATED.  
 CC -I- TISSUE SPECIFICITY: MAMMARY SECRETIONS COLLECTED DURING THE  
 CC NONLACTATING PERIOD.  
 CC -I- SUBCELLULAR LOCATION: EXTRACELLULAR.  
 CC -I- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.  
 DR PIR: A27682; A27682.  
 KW GLYCOPROTEIN.  
 FT NON\_TER 36  
 SQ SEQUENCE 36 AA: 4264 MW: F3935740 CRC32;

Query Match 6.3%; Score 172; DB 4; Length 36;  
 Best Local Similarity 48.6%; Pred. No. 8.27e-13;  
 Matches 17; Conservative 11; Mismatches 7; Indels 0; Gaps 0;

Db 2 klxytswsqyregdskfpaldpflxthviys 36  
 ||:||||:||||:| ||| ||| |||

Qy 2 KLVCYFTNMAQYRQGEARFLPRDLDPISLCTHLIYA 36  
 Search completed: Fri Jun 27 17:00:43 1997  
 Job time : 63 secs.





```

Dh 200 rllldfswlslvdlngswkxvghnspfls-1--pgdtk-sa-yamyywqlvapekll 254
Qy 181 QNldFVNlMAYdFhGSWEKVTGhNSPLKROESGGAASLWDAVQmLQKtPASKLl 240
Dh 255 mqlpbygrtffhllkassqnelragavspaspkyrlkqagflayelccfvrrakkwind 314
Qy 241 LGMpTYGHSFLLASSDPRVCApRtGStGPfLLKESGMlAYEVCSW-KGARKÖRlODp 259
Dh 315 yvpyafkxkewvgyddaisfykafllkrehfgamwclldlfsryfcgtcpfrplvht 374
Qy 300 KVPYlFRDNÖMWGDVDESFKTSYSLKÖKGAGAMWALDLDDFAFGSCNÖGRYPllÖr 359
Dh 375 lnnll 379
Qy 360 LRÖEL 364

```

RESULT	2
ID	R73993; standard; Protein; 668 AA.
AC	R73993.
DT	10-JAN-1996 (first entry)
DE	Hamster oviduct specific glycoprotein.
KW	Hamster oviduct specific glycoprotein; recombinant production;
KW	HOGP.
OS	Cricetulus griseus.
PH	Key
FT	Peptide
FT	/label= sig_peptide
FT	Peptide
FT	/label= mat_peptide
PN	U07107979-A.
PD	25-APR-1995.
PF	15-AUG-1994; 214227.
PR	19-AUG-1993; JP-227881.
PA	(KINOSEI) KINOSEI PEPTIDE KENKYUSHO KK.
DR	WP1: 95-190179/25.
DR	N-PSDB: Q90444.
PT	New DNA encoding an oviduct-specific glycoprotein - useful for
PT	recombinant protein production in high quantities.
PS	Claim 4; Pages 16-18; 22pp; Japanese.
CC	Q90444 encodes R73993 hamster oviduct specific glycoprotein (HOGP).
CC	The cDNA can be used for the commercial recombinant prodn. of
CC	HOGP in high quantities.
SO	Sequence 668 AA;

	Query Match	48.0%	Score 1303	DB 14	Length 668
Qy	Best Local Similarity	48.4%	Pred. No. 2.57e-111		
	Matches 177	Conservative	81	Mismatches 99	Indels 9
Db	20 klvcyfcfwnahsrpypasip1prdiidpfc1thl1afafasmsnqivam1dck1lypefn	79			
Qy	2 klvcyfcfwnahsrpypasip1prdiidpfc1thl1afafasmsnqivam1dck1lypefn	60			
Db	80 klkernalkcl1slygvwmfstrfctm1stclslarek1figsvsflrtfthgfdgld1fily	139			
Qy	61 glkknmpklktll1aigcwnfctokfctdmv1aannrqcfvnsa1rfekrksyfdgld1mex	120			
Db	140 pglisgplndtwfn1f1lee1qfefekeall1tqrpl1saav-sg1py1l1c1s1dyvhl	198			
Qy	121 psgosspavdxkerf1tllv1o1an1f0e1a1n1sg1er1ul1sa1v1a1g1t1y1-d1ge1y1edk1	179			
Db	199 gtridf1nvl1sydl1hgsweks1tghnsp1s1f1pedp1ksa--fann--y--wrn1gapack1	253			
Qy	180 aon1ndfvm1lma1vdfgsm1ex1tghn1sp1l1krn1qo1ess1ga1as1l1vda1va1oq1m1l1o1ng1t1p1sk1	239			
Db	254 lmgfpa1yrtf1h1l1reskn1l1qaasm1g1pas1p1k1y1k1ag1f1ay1evcs1f1q1raek1m1d1h	313			
Qy	240 ilgm1tyr1s1et1ll1a1ss1s1p1r1vga1a1t1g1s1g1p1t1k1eg1m1l1ay1evcs1w--kg1a1t1k1o1r1o1d	298			
Db	314 q1y1p1ay1r1g1k1ew1y1d1v1s1f1sy1kam1f1k1k1h1fg1an1v1t1l1md1d1v1r1y1t1c1g1n1p1f1v1h	373			
Qy	299 kv1p1e1f1d1n1o1w1g1d1v1e1s1k1t1k1v1s1y1k1o1g1l1g1a1m1w1a1l1d1d1f1a1g1s1c1n1o1g1r1p1l1o	358			

```
Db      374 11n11 379
        | : |
QY      359 TLRL 364
```

ID	RESULT	3
AC	R73992; standard; protein; 718 AA.	
DT	10-JAN-1996 (first entry)	
DE	Murine oviduct specific glycoprotein.	
KW	Murine oviduct specific glycoprotein; recombinant production;	
KW	MOGP.	
OS	Mus musculus.	
EH	key	Location/Qualifiers
FT	Peptide	1..18
FT	/label= sig_peptide	
FT	Peptide	19..718
FT	/label= mat_peptide	
PN	J07107979-A.	
PD	25-APR-1995.	
PF	15-AUG-1994; 214227.	
PR	19-AUG-1993; JP-227881.	
PA	(KINO-) KINOSEI PEPTIDE KENKYUSHO KK.	
DR	WPI; 95-190179/25.	
DR	N-PSDB: Q90443.	
FT	New DNA encoding an oviduct-specific glycoprotein - useful for	
FT	recombinant protein production in high quantities.	
PS	Claim 4; Pages 11-14; 22p; Japanese.	
CC	Q90443 encodes R73992 murine oviduct specific glycoprotein (MOGP).	
CC	The cDNA can be used for the commercial recombinant prodn. of	
CC	MOGP in high quantities.	
SO	Sequence 718 AA;	

	Query Match	48.0%	Score 1303	DB 14	Length 718
	Match Local Similarity	49.9%	Pred. No. 2.57e-111		
	Matches 182	Conservative	76	Mismatches 100	Indels 7
					Gaps 4
Db	20 klvccftmavnsrpgpasimpdhidpfcitlllifaasmsnqivaknlgdenvltypefn	79			
OY	2 klvcftmwaqyrogearfepkrddpslcthlilafragmtthqlsttewnder-lyqefn	60			
Db	80 klkernrelkllslgwnfgtsrftamslstlanrekfidsvslfzrhngfdagldlfeily	139			
OY	61 glkknnpkrlktllamagcnfnfqtokftdmvvaannrqfvensalrflkrkyspdgdldmex	120			
Db	140 pqlrgspdhwnflfielqgaferealltqhprlllsaavgsipslintsydaillg	199			
OY	121 pgsosppavdkerfttllqodlanpfoqeaqrsgkerlllsaaVPAGQYVAGGEVDKIA	180			
Db	200 rrlrdinlsydlngswkftgfnspslfslpedsksa--y---ammywrlgypackll	254			
OY	161 qnldevnlmaadpfnoswkwrtvgshnslpkroeeGGAASLNVDAAYQOMLOKGRPASKLI	240			
Db	255 mgfpygrnfyllkesknlglsaqpaspgkykkgagflayvevcsfvgrakkhwidyq	314			
OY	241 lgmprtgssftlassdsdrvcapaprtgscstpprfkkgcmllryevcsn-kkatrQRIQDQ	299			
Db	315 yvpyafkewlgyddctsfyskamyvkrhehfggamwvltlmdadvrgtfcgnpfpvlh	374			
OY	300 kvpyfyrnqmwgfpdvesfktkysylkQKOLGGMAMWALDLDPAFGSCNQSGKYPILQI	359			
Db	375 lneell 379				
OY	360 lRQEL 364				
RESULT	4				
ID	P81342 standard; protein; 321 AA.				
AC	P81342;				
DT	19-OCT-1990 (first entry)				
DE	Polypeptide involved in protective mechanisms				
KW	Immune response; cell growth.				
PN	J63032898-A.				

PD 30-JAN-1988. 167518.  
 PF 16-JUL-1986; JP-167518.  
 PR 16-JUL-1986; JP-167518.  
 (DAIN) Dainippon Pharm KK.  
 PA WPI; 88-068419/10.  
 DR P-SDB; P81342.  
 PT New polypeptide and DNA encoding it - related to protective mechanisms such as immune response etc.  
 PT Disclosure: p; Japanese.  
 PS This polypeptide is involved in protective mechanisms such as immune response, cell growth and activation of protective functions.  
 CC Sequence 321 AA;

ID	RESULT	5	standard; Protein: 371 AA.
AC	R70025;		
DT	29-SEP-1995		(first entry)
DE	Tobacco chitinase encoded by tobacco Cluster-A cDNA clone CA-3.		
KM	Tobacco; chitinase; antifungal; fungicide; Cluster-A.		
OS	Nicotiana tabacum.		
FH	Key	Location/Qualifiers	
FT	Peptide	1..18	
FT	/label= signal peptide		
FT	misc_difference 22		
FT	/note= "not present in genomic clone"		
FT	Misc_difference 182..195		
FT	/note= "Determined by protein sequencing"		
FT	Misc_difference 216..241		
FT	/note= "Determined by protein sequencing"		
PN	EP-639642-A.		
PD	22-FEB-1995.		
PF	17-AUG-1993; 202425.		
PR	17-AUG-1993; EP-202425.		
PA	(MOGE-) MOGEN INT NV.		
PA	(OYLE-) RIKSDUIT LEIDEN.		
PI	Altotheker-de Groot M, Bol JF, Cornelissen BJC, Linthorst HJM;		
PI	Melchers IS, Ponstein AS, Sela-bourlaga MB;		
DR	WPI: 95-083454/12.		
DR	N-PSDB; 082973.		
PT	New plant protein having endo-chitinase activity - used in		
PT	antifungal compms. and to develop transformed plants which are		
PT	less susceptible to fungal infection.		
PS	Claim 3; Page 19-20; 43pp; English.		
CC	Screening of a lambda ZAP cDNA library of TMV-infected Sansun NN		
CC	tobacco plants with a probe derived from P80840 (a partial Cluster-A		
CC	cDNA clone) resulted in the isolated of 11 positively hybridising		
CC	clones. Analysis revealed that all were identical and corresp. to		
CC	Cluster-A cDNA. The nt sequence of clone CA-3 is given in 082973/		
CC	R70025. This clone lacks 7 codons of the 5' part of the ORF. The		
CC	cDNA clone was completed by performing a PCR reaction on clone CA-3		
CC	with primer T7 (082794) and primer P1 (082975). A partial AA		
CC	sequence of the protein isolated from TMV-infected tobacco leaves		

CC was obtd. and found to be almost identical to the AA sequence  
CC deduced from the cDNA.  
SQ Sequence 371 AA;

Query Match	22.0%	Score 597	DB 13	Length 371
Best Local Similarity	33.0%	Pred. No. 1,076	43	
Matches 107	Conservative	86	Mismatches 108	Indels 23
			Gaps 18	
Db	36 nidslftfhcfcafdlnpsnqllispends-fsgfscvtgkrnpvktfnslagqr 94			
Oy	23 KDLDPSSLCHLYAFAGCT--NHOLDSTENDELVEEPFG-LKKMKPKLTLLAI-GCW 78			
Db	95 a-dttayigimarqpnstrksfidsisrlarqfghlglidweyp--l-s-atdmtnlglil 149			
Oy	79 NFGTQKFMDMATATNNRQTFEWSAIRFLRKISFPGLDLMDEYPCQSQSPAVDKERFTTLV 138			
Db	150 newrtalimernngiraallltaavsyspvrnglnyvesvarnlmnlmaydfygnw 209			
Oy	139 QDLANAFQEQEAGTGSKERLLLSAAPVPAQGYTVDAEGYEDKIAQNLMDLMAAYDFEHS-W 197			
Db	210 spsqtnshaqfddpvnh--lsgs---ag-lnaavagavpckklvlg4pfgyawrlvnpn 263			
Oy	198 EKV-TGHSPIKXROEBEGGAALSLNDAAVQWQWQKTPASKLLTGMPYIGRSTLLSSS 256			
Db	264 ihdlrapaagksnvga-vddgsm-tynrirdyivqtrattynatlvgdgcysgsnwlslsy 321			
Oy	257 DTRVGAIPATGSGTGPFPFKTEGGLMAYEVCWKGATKQR-IQDQK-VP-YIFRDNQWVF 313			
Db	322 ddtgsvrnknvnyvgqrglgyfaw 345			
Oy	314 DDVESFKTKVSYLKOKGLGAMW 337			

RESULT	6	
ID	R70029	standard; Protein: 377 AA.
AC	R70029;	
DE	29-SEP-1995	(first entry)
DD	Tobacco cluster-A protein encoded by genomic clone.	
DE	Tobacco cluster-A protein encoded by genomic clone.	
KW	Tobacco; chitinase; antifungal; fungicide; Cluster-A.	
OS	Nicotiana tabacum.	
FS	Key	Location/Qualifiers
FT	Peptide	1..25
FT	/label= signal peptide	
FT	Misc.difference 28..29	
FT	/note="cDNA encodes additional Ile here"	
FT	Misc.difference 188..201	
FT	/label= Determined by sequencing the protein	
FT	Misc.difference 222..247	
FT	/note= "Determined by sequencing the protein"	
PN	EP-639642-A.	
PD	22-FEB-1995.	
PF	17-AUG-1993; 202425.	
PR	17-AUG-1993; EP-202425.	
PA	(MOGE-) MOGEN INT NV.	
PA	(UYLE-) RIJSDIJK LEIDEN.	
PI	Apotheker-de Groot W, Bol JF, Cornelissen BJC, Linthorst HJM;	
PI	Melchers LS, Ponsstein AS, Sela-buurlage MB;	
DR	WPI: 95-083454/12.	
DR	N-PSDB: 082976.	
PT	New plant protein having endo-chitinase activity - used in	
PT	antifungal compsns. and to develop transformed plants which are	
PT	less susceptible to fungal infection.	
PS	Claim 3; Page 25-28; 43pp; English.	
CC	Screening of a lambda Zap cDNA library of TMV-infected Samsum NN	
CC	tobacco plants with a probe derived from PRO340 (a partial Cluster-A	
CC	cDNA clone) resulted in the isolated of 11 positively hybridising	
CC	clones. Analysis revealed that all were identical and corresp. to	
CC	Cluster-A cDNA. The sequence of cDNA clone CA-3 is given in 082973/	
CC	R70029. A genomic library of N. tabacum was screened using the	
CC	Cluster-A cDNA. Insert of clone CA-3 as a probe (see 082977, 082978).	
CC	The complete nt. sequence of Cluster-A cDNA including the deduced	
CC	primary structure of the Cluster-A protein the 5' and 3' UTR regions	
CC	of the gene are shown in 082976/870029. Comparison of the cDNA	



OY 301 VPIYF-RD-NQWVGFDVSEFRTKVSYLKOKGLGAMW 337

RESULT 9  
ID R33072 standard; Protein: 423 AA.  
AC R33072;  
DT 30-JUN-1993 (first entry)  
DE Aphanocladium album pre-pro-endochitinase.  
KM filamentous fungus; pathogen resistance.  
OS Aphanocladium album.  
PI Blaiseau P, Legoux R, Leguay J, Schneider M;  
DR WPI: 93-078700/10.  
PT Recombinant DNA encoding fungal endo:chitinase - useful for  
PS protecting plants against fungi, insects, etc.  
CC Claim 1: Fig 1: 75bp. French.  
CC Clone CH3C was isolated from a cDNA bank prepared from A. album  
CC mRNA by screening with anti-chitinase antiserum. Three probes were  
CC designed based on the CH3C sequence, corresponding to the start,  
CC middle and downstream regions of the coding sequence. The probes  
CC were used to isolate the full-length coding sequence of A. album  
CC endochitinase from an A. album cDNA library in pTZ19R. A 1.6kb  
CC fragment was identified in several positive clones. The sequence  
CC of the fragment contained an ORF encoding a 423 amino acid protein.  
CC The mature endochitinase can be used to control pathogens  
CC such as fungi, bacteria, arthropods and nematodes on plants.  
SQ Sequence 423 AA:

Query Match 21.2%; Score 576; DB 6; Length 423;  
Best Local Similarity 36.2%; Pred. No. 1.01e-41;  
Matches 101; Conservative 65; Mismatches 95; Indels 18; Gaps 16;

DB 115 LKQGRNMKVMISIGVWST-NFPAASSAATKTFAGSAYGFMKDWGFGDIDWEYP 173  
OY 62 LKKNPKLTKTLAIGMNGFTQKFTDMVATANNROTFFVNSAIRFLRKYSFSDGLDWEYP 121

DB 174 -adatqagmvlilqavrseldsyaay-akg-hhflislaapadpnykikfa-e-1g 228  
OY 122 GSGGSADVAVKERTTLVQDLANAFQEAQTSGERLLLSAAVPAG-QTYVDAGYEVDKIA 180

DB 229 kvldylnmaydyagswnytghdaniyanpnunatp-yntddavqayingvpankiv 287  
OY 181 QNLDPEVNLMAVDFHGSMEVYTGNSPLRYKROESGAASLNDAVQOVLQKTPASKLI 240

DB 288 lmplyvrsifg-qt--eg-1gkpyng1gs-gsw--englwdykalpk-agatvkcdttak 339  
OY 241 LQMPTVGRSFTLASSSDTRVGAAPATGSGTPFGPTKREGMLAYEVCSWKGATKQRIQDOK 300

DB 340 gcysydpstckelstfdtpamistkxswlkgk1gsgmfw 378  
OY 301 VPIYF-RD-NQWVGFDVSEFRTKVSYLKOKGLGAMW 337

RESULT 10  
ID R33068 standard; Protein: 389 AA.  
AC R33068;  
DT 30-JUN-1993 (first entry)  
DE Aphanocladium album mature endochitinase.

KM filamentous fungus; pathogen resistance.  
OS Aphanocladium album.  
PN EP-531218-A.  
PD 10-MAR-1993.  
PF 04-SEP-1992; 402414.  
PR 06-SEP-1991; FR-011072.  
PA (SNFI) ELF SANOFI.  
PI Blaiseau P, Legoux R, Leguay J, Schneider M;  
DR WPI: 93-078700/10.  
DR N-PSDB: Q37704.  
PT Recombinant DNA encoding fungal endo:chitinase - useful for  
PS protecting plants against fungi, insects, etc.  
CC Claim 22; Page 60-61; 75bp. French.  
CC This mature protein from the filamentous fungus Aphanocladium  
CC album has endochitinase activity and is used to control pathogens  
CC such as fungi, bacteria, arthropods and nematodes on plants.  
CC Recombinant DNA comprising the endochitinase coding sequence is  
CC specifically intended for transforming Nicotiana tabacum, Helianthus  
CC annuus and Brassica napus to pathogen resistance.  
SQ Sequence 389 AA:

Query Match 20.9%; Score 569; DB 6; Length 389;  
Best Local Similarity 35.8%; Pred. No. 4.58e-41;  
Matches 100; Conservative 66; Mismatches 95; Indels 18; Gaps 16;

DB 81 LKQGRNMKVMISIGVWST-NFPAASSAATKTFAGSAYGFMKDWGFGDIDWEYP 139  
OY 62 LKKNPKLTKTLAIGMNGFTQKFTDMVATANNROTFFVNSAIRFLRKYSFSDGLDWEYP 121

DB 140 -adatqagmvlilqavrseldsyaay-akg-hhflislaapadpnykikfa-e-1g 194  
OY 122 GSGGSADVAVKERTTLVQDLANAFQEAQTSGERLLLSAAVPAG-QTYVDAGYEVDKIA 180

DB 195 kvldylnmaydyagswnytghdaniyanpnunatp-yntddavqayingvpankiv 253  
OY 181 QNLDPEVNLMAVDFHGSMEVYTGNSPLRYKROESGAASLNDAVQOVLQKTPASKLI 240

DB 254 lmplyvrsifg-qt--eg-1gkpyng1gs-gsw--englwdykalpk-agatvkcdttak 305  
OY 241 LQMPTVGRSFTLASSSDTRVGAAPATGSGTPFGPTKREGMLAYEVCSWKGATKQRIQDOK 300

DB 306 gcysydpstckelstfdtpamistkxswlkgk1gsgmfw 344  
OY 301 VPIYF-RD-NQWVGFDVSEFRTKVSYLKOKGLGAMW 337

RESULT 11  
ID W02159 standard; Protein: 866 AA.  
AC W02159;  
DT 14-JAN-1997 (first entry)  
DE Soluble chitinase.  
KM Periplasmic chitodextrinase; periplasmic beta-N-acetylglucosaminidase;  
KW beta-N-acetylglucosaminidase; chitin; oligosaccharide; catabolic;  
OS catabolism.  
OS Vibrio furnissii.  
PN W09625424-A1.  
PD 22-AUG-1996.  
PF 13-FEB-1996; U02332.  
PR 13-FEB-1995; US-386727.  
PA (UYIO) UNIV JOHNS HOPKINS.  
PI Bassler B, Chitliaru E, Keyhani N, Roseman S, Rowe C;  
DR WPI: 96-39335/39.  
DR N-PSDB: T36390.  
PT Chitin biosynthetic enzymes and I, exo I and exo II - are  
PT periplasmic chitodextrinase(s), periplasmic beta-GlcNAcidase(s) and  
PT aryl beta-N-acetylglucosaminidase(s), respectively  
Example 4; Page 79-82; 101pp; English.  
CC Periplasmic chitodextrinase (W02156), periplasmic  
CC beta-N-acetylglucosaminidase (W02157) and aryl  
CC beta-N-acetylglucosaminidase (W02158) can be used to produce chitin  
CC oligosaccharides with the structure (GlcNAc)n where n is 2 or



[illegible]

RESULT 14  
ID R32547 standard; Protein: 238 AA.  
AC R32547;  
DT 17-JUN-1993 (first entry)  
DE Hsnpv IE-1 gene protein (partial sequence).  
KW Heliothis armigera nuclear polyhedrosis virus; Ha; NPV;  
KW Insecticide; immediate early-1.  
OS Heliothis armigera nuclear polyhedrosis virus.  
PN W09303144-A.  
PD 18-FEB-1993.  
PF 05-AUG-1992; AU0413.  
PR 05-AUG-1991; AU-007576.  
PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
PI Christian PD, Dail DJ, Gordon KHT, Hanzlik TN, Srisatchana A.  
DR WPI: 93-075601/09.  
NR PDB: Q35882.  
DT Insect virus with reduced capacity to occlude viral particles - used  
PT for controlling proliferation of insect pests without horizontal  
PT transmission  
PS Disclosure: Fig 2: 51pp: English.  
CC The sequence is that encoded by the Heliothis armigera (Ha)  
CC nuclear polyhedrosis virus (NPV) immediate early-1 (IE-1) gene  
CC (partial sequence).  
SQ Sequence 238 AA;

RESULT	ID	Accession	Description
15	AC	W02156	standard; Protein; 1046 AA.
16	AC	W02156	
17	DT	14-JAN-1997	(first entry)
18	DE		Periplasmic chitodextrinase.
19	KW		Periplasmic chitodextrinase; periplasmic Beta-N-acetylglucosaminidase
20	KW		Beta-N-acetylglucosamidase; chitin; oligosaccharide; catabolic;
21	KW		catabolism.
22	OS	Vbtrp01_furniss11.	
23	PN	M09625424-A1.	
24	PD	22-AUG-1996.	
25	PF	13-FEB-1996.	U02332.
26	PR	13-FEB-1995.	US-386727.
27	PA	(UUYO ) UNIV JOHNS HOPKINS.	
28	PI	Baszler B, Chitlaru E, Keyhani N, Roseman S, Rowe C;	
29	PI	Yu C,	
30	DR	WPI: 96-393335/39.	
31	DR	N-PSDB: T36387.	
32	PT		Chitin blosynthetic enzymes end I, exo I and exo II - are

PT periplasmic chitin:deacetylase(s), periplasmic beta-glucanase(s) and  
PT aryl beta-N-acetylglucosaminidase(s), respectively  
PS Claim 2; Page 68-71; 101pp: English.  
CC Periplasmic chitin:deacetylase (W02156), periplasmic  
CC Beta-N-acetylglucosaminidase (W02157), and aryl  
CC Beta-N-acetylglucosaminidase (W02158) can be used to produce chitin  
CC oligosaccharides with the structure (GlcNAc)<sub>n</sub> where n is 2 or  
CC higher, by contacting them with soluble chitin. The enzymes are  
CC encoded by the genes *exl1*, *exl2* and *exl3* respectively. They are  
CC all genes involved in the catabolic pathway of chitin.  
SQ Sequence 1046 AA;

Search completed: Fri Jun 27 17:10:12 1997  
Job time : 102 secs.





\*\*\*\*\*  
 WISE (TM)  
 \*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
 Distribution rights by Intelligenetics, Inc.

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
 Run on: Fri Jun 27 17:04:42 1997; Maspar time 17.76 Seconds  
 Tabular output not generated.  
 598.666 Million cell updates/sec

Title: >US-08-663-618A-15  
 Description: (1-373) from US08663618A.pep  
 Perfect Score: 2717  
 Sequence: 1 AKLYCYFTNMAQYRQGEARF.....YPLIQTLRQELSLPYLSSGT 373

Scoring table:  
 PAM 150  
 Gap 11

Searched: 89912 segs, 28507787 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database:

pir50  
 1:ann1 2:ann2 3:ann3 4:ann4 5:unann1 6:unann2 7:unann3  
 8:unann4 9:unann5 10:unann6 11:unann7 12:unann8  
 13:unann9 14:unann10 15:unann16:unrev

Statistics: Mean 47.342; Variance 104.126; scale 0.455

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	1488	54.8	383	A49562	cartilage glycoprotein	1.46e-250
2	1428	52.6	383	S51327	heparin-binding glycoprotein	5.60e-237
3	1417	52.2	381	S61550	BRP39 protein - mouse	5.60e-237
4	1417	52.2	381	S61550	BRP39 protein - mouse	5.60e-237
5	1345	49.5	399	S27879	secretory protein YM	5.55e-220
6	1328	48.9	537	S57197	oviduct-specific glycoprotein	4.49e-218
7	1318	48.5	539	I46470	estrogen dependent oviduct-specific glycoprotein	1.28e-191
8	1179	43.4	483	A53918	chitinase (EC 3.2.1.1)	5.49e-174
9	1086	40.0	504	A38221	chitinase (EC 3.2.1.1)	1.44e-160
10	1015	37.4	554	A56596	chitinase (EC 3.2.1.1)	3.36e-82
11	698	25.7	699	A38368	chitinase (EC 3.2.1.1)	5.18e-80
12	594	21.9	378	S51591	chitinase (EC 3.2.1.1)	6.42e-79
13	582	21.4	424	S47133	chitinase (EC 3.2.1.1)	3.43e-76
14	576	21.2	423	S51369	chitinase (EC 3.2.1.1)	1.80e-73
15	561	20.6	423	S51369	chitinase (EC 3.2.1.1)	1.80e-73
16	546	20.1	427	UC4365	chitinase precursor	4.21e-48
17	446	16.4	561	A52090	chitinase (EC 3.2.1.1)	7.52e-38
18	404	14.9	561	A52090	chitinase (EC 3.2.1.1)	1.12e-37
19	344	12.7	799	PC4106	chitinase (EC 3.2.1.1)	1.12e-37
20	345	12.7	820	A40653	chitinase (EC 3.2.1.1)	1.12e-37
21	344	12.7	820	A40653	chitinase (EC 3.2.1.1)	1.12e-37

RESULT ENTRY	1	22	337	12.4	499	9	S04856	chitinase (EC 3.2.1.1)
TITLE	A49562	314	11.6	511	11	S61166	chitinase (EC 3.2.1.1)	
ALTERNATE_NAMES	39K synovial protein	24	10.6	610	10	UC0573	chitinase (EC 3.2.1.1)	
ORGANISM	Homo sapiens	25	9.8	452	12	UC4038	47K glycoprotein	
DATE	03-May-1996	26	9.7	1146	11	S07915	chitinase - yeast	
ACCESSIONS	A49562; S10677; A33162	27	9.1	413	5	UC2135	chitinase (EC 3.2.1.1)	
REFERENCE	A49562; S10677; A33162	28	8.4	124	5	S57715	chitinase (EC 3.2.1.1)	
Journal	Hakala, B. E.; White, C.; Recklies, A. D.	29	6.3	36	14	A27682	39K whey protein - b	
#title	Human chitinase gp-39, a major secretory product of articular chondrocytes and synovial cells, is a mammalian member of a chitinase protein family.	30	6.3	83	13	A37954	estradiol-stimulated	
Accession	A49562	31	5.0	29	14	S57204	oviduct-specific gly	
Status	preliminary	32	5.0	29	14	S55693	oviduct-specific gly	
Molecule-type	mRNA	33	4.5	597	5	S33849	chitinase (EC 3.2.1.1)	
Residues	1-383	34	4.5	597	5	S32039	chitinase (EC 3.2.1.1)	
Cross-references	GB:M80927	35	4.1	642	16	S37623	homologues with orf	
Authors	Nytkos, P.; Golds, E. E.	36	4.0	488	10	A41961	chitinase (EC 3.2.1.1)	
Journal	Biochem. J. (1990) 265:265-268	37	3.8	305	9	S28440	probable coproporph	
Title	Human synovial cells secrete a 39 kDa protein similar to a bovine mammary protein expressed during the non-lactating period.	38	3.8	422	4	ZWEC	hypothetical protein	
Accession	A49562	39	3.7	364	11	S45925	keratin, 56K type II	
Journal	Hakala, B. E.; White, C.; Recklies, A. D.	40	3.7	562	3	KRUEB	probable membrane pr	
#title	Human chitinase gp-39, a major secretory product of articular chondrocytes and synovial cells, is a mammalian member of a chitinase protein family.	41	3.7	662	11	S61193	serine proteinase (E	
Accession	A49562	42	3.7	1045	2	SUSEM	alpha-glucosidase (E	
Status	preliminary	43	3.6	1070	11	S19686	streptococcal pyroge	
Molecule-type	mRNA	44	3.6	398	5	A37768	glucan 1,4-alpha-glu	
Residues	1-383	45	3.6	958	1	JN0102	3.26e+00	
Cross-references	GB:M80927							

#### ALIGNMENTS

Query Match 54.8%; Score 1488; DB 13; Length 383;  
 Best Local Similarity 53.3%; Pred. No. 1.46e-250;  
 Matches 195; Conservative 83; Mismatches 79; Indels 9; Gaps 5;

Db 23 klvcytsvsgyregqscfpaldlftclthlysfanlsndhltwewdvlygmnt 82  
 |||||:|||||:|:| |||||:||||:|:| ||||| ||| :|  
 2 klvcyftnmaqrqgarflpdlpslcthllyafagwtmhnqlstewndetlyqefng 61  
 |||||:|||||:|:| |||||:||||:|:| ||||| ||| :|  
 Db 83 lknrpnltkllsvsgvsgsfktsantqsttrflksvppflrtfngf6glaiawlyp 142  
 |||||:|||||:|:| |||||:||||:|:| ||||| ||| :|

```

Oy 62 LKKNPKLKTLLAIGWMEFGTQKFTDMVATANNROTFFVNSAIRFLRKXSFSDLDIMWEXR 121
Db 143 grf-----dkchfttlkemkaeflkeag-pqkqglllisaalsagkvtldssydaiklsq 196
Oy 122 GSQGSFPAVDKEFTTLVODLANAFQOEAOQTSKERLLLSAAVPAQOTYVDAGYEDKTAQ 181
Db 197 hldfslmtcydfhagwrgtctghspilfrgqedasprfnsctdyagymrlrlgapasklvm 256
Oy 182 NLDFVNLMAYDFHGSWEKVTGHNSPLYKROESGAAASLNDVANAQOMLQKGTSPASKLIL 241
Db 257 gipctgrfctlaas-ctvgapapispipgrfctkeagctlayeylclfrgatvhtlqg 315
Oy 242 GMPYGRSFITLASSSDTRVGAAPATGSGTPGPPTKEGMLAYEVCWSM-KGATKQRIQOK 300
Db 316 VPYATKQNGWYGDGGSVKSQVYJLKDRLQAGAMWALDIDFGSGFCGQDLRFPLNA 375
Oy 301 VPYTRDNQMWGFDDVESFKTKVSTLKQKGLGAGAWMALDDDFAGFCSCNOG-RYPLLIOT 359
Db 376 lkdala 381
Oy 360 LROELS 365

RESULT 2
ENTRY S51327 #type complete
TITLE heparin-binding glycoprotein 38k - pig
ORGANISM #formal_name Sus scrofa domestica #common_name domestic pig
DATE 07-May-1995 #sequence_revision 07-May-1995 #text_change
ACCESSIONS S51327
REFERENCE S51327
#authors Shackelton, L.M.; Mann, D.M.; Mills, A.J.T.
#description Submitted to the EMBL Data Library, January 1995
#description Identification of a 38kda heparin-binding glycoprotein
#description (gp38k) in differentiating vascular smooth muscle cells as
#description a member of a group of proteins associated with tissue
#description remodeling.
#accession S51327
#status preliminary
#residues 1-383 #label SHN
#cross-references EMBL:247803
SUMMARY #length 383 #molecular_weight 42443 #checksum 2907

Query Match 52.6%; Score 1428; DB 16; Length 383;
Best Local Similarity 52.2%; Pred. No. 4,41e-239;
Matches 191; Conservative 81; Mismatches 85; Indels 9; Gaps 5;

Db 23 klvcytswgryregdscfpdaipfctihlysfanisdmlclwmdvcltydltnt 82
Oy 2 KLVCYFTWMAQYRGGEARFLPKDLPSCITHLIYAFAGMTNHLSTEMNDETLVQEFN 61
Db 83 lknrpnklktllsvgvgfsggrfctksnctgrrtftksvppflrtfngfgldlawnsp 142
Oy 62 LKKNPKLKTLLAIGWMEFGTQKFTDMVATANNROTFFVNSAIRFLRKXSFSDLDIMWEXR 121
Db 143 grf-----dkchfttlkemkaeflkeag-pqkqglllisaalsagkvtldssydaiklsq 196
Oy 122 GSQGSFPAVDKEFTTLVODLANAFQOEAOQTSKERLLLSAAVPAQOTYVDAGYEDKTAQ 181
Db 197 hldfslmtcydfhagwrgtctghspilfrgqedasprfnsctdyagymrlrlgapasklvm 256
Oy 182 NLDFVNLMAYDFHGSWEKVTGHNSPLYKROESGAAASLNDVANAQOMLQKGTSPASKLIL 241
Db 257 gipctgrfctlaas-ctvgapapispipgrfctkeagctlayeylclfrgatvhtlqg 315
Oy 242 GMPYGRSFITLASSSDTRVGAAPATGSGTPGPPTKEGMLAYEVCWSM-KGATKQRIQOK 300
Db 316 VPYATKQNGWYGDGGSVKSQVYJLKDRLQAGAMWALDIDFGSGFCGQDLRFPLNA 375
Oy 301 VPYTRDNQMWGFDDVESFKTKVSTLKQKGLGAGAWMALDDDFAGFCSCNOG-RYPLLIOT 359
Db 376 lkdala 381
```

```

Oy 360 LROELS 365

RESULT 3
ENTRY S61550 #type complete
TITLE BRP39 protein - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change
ACCESSIONS S61550
REFERENCE S61550
#authors Morrison, B.W.
#description Submitted to the EMBL Data Library, November 1995
#accession S61550
#status preliminary
#residues 1-381 #label MOR
#cross-references EMBL:X93035
SUMMARY #length 381 #molecular_weight 43001 #checksum 5107

Query Match 52.2%; Score 1417; DB 14; Length 381;
Best Local Similarity 50.7%; Pred. No. 5,60e-237;
Matches 186; Conservative 84; Mismatches 86; Indels 11; Gaps 9;

Db 23 klvcytswgryregdscfpdaipfctihlysfanisdmlclwmdvcltydltnt 82
Oy 2 KLVCYFTWMAQYRGGEARFLPKDLPSCITHLIYAFAGMT-NHOLSTTEMNDETLVQEFN 60
Db 83 klctnrlktllsvgvgkfgkrfseiasnterrtafvsvapflrsyfgdglawly 142
Oy 61 GLKKNPKLKTLLAIGWMEFGTQKFTDMVATANNROTFFVNSAIRFLRKXSFSDLDIMWEXR 120
Db 143 p-r-lr--dkyftslkleinaeftevg-pyreklllisaalsagkvvaldtydldia 196
Oy 121 PGSQGSFPAVDKEFTTLVODLANAFQOEAOQTSKERLLLSAAVPAQOTYVDAGYEDKTA 180
Db 197 qhldfslmtcydfhagwrgtctghspilfrgqedasprfnsctdyagymrlrlgapaskl 256
Oy 181 QNDFVNLMAYDFHGSWEKVTGHNSPLYKROESGAAASLNDVANAQOMLQKGTSPASKLIL 240
Db 257 mgitftgkscftlaas-englqapispigpgrfctkeagctlayeylclfrgaevhrisne 315
Oy 241 LGMPYGRSFITLASSSDTRVGAAPATGSGTPGPPTKEGMLAYEVCWSM-KGATKQRIQOK 299
Db 316 kvpatkngqvgvghesvknkgfllkekklagawwalddfgg-tcqpkefip1tn 374
Oy 300 KVPYTRDNQMWGFDDVESFKTKVSTLKQKGLGAGAWMALDDDFAGFCSCNOGRY-PLIQ 358
Db 375 alkdala 381
Oy 359 TLROELS 365

RESULT 4
ENTRY I48271 #type complete
TITLE BRP39 protein - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change
ACCESSIONS I48271
REFERENCE I48271
#authors Morrison, B.W.; Leder, P.
#journal Oncogene (1994) 9:3417-3426
#title neu and ras initiate murine mammary tumors that share genetic
#title markers generally absent in c-myc and int-2-initiated
#title tumors.
#cross-references MUID:95060797
#accession I48271
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-381 #label RES
#cross-references EMBL:X93035; NID:g1085065; CDS_PID:g1085066
```



	Query Match	43.4%;	Score 1179;	DB 12;	Length 483;
	Best Local Similarity	42.5%;	Pred. No. 1,28e-191;		
	Matches 158;	Conservative	98;	Mismatches 99;	Indels 17; Gaps 11,
Db	23	kvvvyfgawsvyrrgnkfdingldpblchlllysfygvngkdvkvlidpmsldipnldgf	82		
	1				
Qy	2	KLVEYFNMAQYRGGEARFLPKDIDPSLCHLLLYAFAGMNHOLSTTE-WND--ETL--Y	56		
Db	83	gkftslrknpvsvklnwavgwagvsrpfasqmasdgaatreafanvkvflqgyfgdgd	142		
	1				
Qy	57	QEFVGLKKNMKNKLTLLAIGWMNGCTOKFTDMATANNRQTFVNSAIRFLRKTSFSDLD	116		
Db	143	dweydaqrggsapadvkmvklckalkkafqg-h---d-yilsaavaapetsasksyd1	195		
	1				
Qy	117	DMEYPGSGSFAVDEKERFTTLVLDLANAFOEQATSGKERILLSAAPACQTYVDAGYEV	176		
Db	196	aemsgyldfinlmtydthgpmddgtlqmhappsasahsgneiklnvkaevkylqgvvk	255		
	1				
Qy	177	DKIQNIDFVWLMKYDRHGSWEKTEGHNSPLYKRQESSGAASLNDAAVQWLOKQTPA	236		
Db	256	eklvvgyaygkfstlaspnknlgapavsgatgapytgcnglllynelcemqkagdev	315		
	1				
Qy	237	SKLLIGMPTGRSPTLLASSSDTRRGARATSGSPGPTKRGGMILAYEVSW-KGATKQR	295		
Db	316	vqdmekvpyavkngvnsfddlaialakxagqflkqegllgagmwslctddfkgl-cge-k	373		
	1				
Qy	296	IQD-QK-VPIYFRDNQWGVDFDVESEFKTKYVKOKGLGAMWALDIDDFAGFSCNQR	353		
Db	374	ypvikaalnsyvj	385		
	1				
Qy	354	YPLQTLRQELS	365		
RESULT	9				
ENTRY	A38221	#type complete			
TITLE	chitinase (EC 3.2.1.14) Mf1 - nematode (Brugia malayi)				
ORGANISM	formal name Brugia malayi				
DATE	04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 03-May-1996				
ACCESSIONS	A38221				
REFERENCE	A38221				
#authors	Fuhrman, J.A.; Lane, W.S.; Smith, R.F.; Plessens, W.F.; Pearl, F.B.				
#journal	Proc. Natl. Acad. Sci. U.S.A. (1992) 89:1548-1552				
#title	Transmission-blocking antibodies recognize microfilarial chitinase in Brugian lymphatic filariasis.				
#cross-references	MOTID:92179220				
#accession	A38221				
##status	preliminary; not compared with conceptual translation				
##molecule_type	nucleic acid: protein				
##residues	1-504 ##label FUH				
##cross-references	NCBI:P:85345				
##note	sequence extracted from NCBI backbone				
KEYWORDS	glycosidase; hydrolase				
SUMMARY	#length 504 #molecular-weight 55971 #checksum 7243				
Query Match	40.0%;	Score 1086;	DB 12;	Length 504;	
Best Local Similarity	42.8%;	Pred. No. 5.49e-174;			
Matches 158;	Conservative	94;	Mismatches 102;	Indels 15; Gaps 11,	
Db	27	cytcnwaqyrdgqekflipnigpnlcjhllyafakvdelgskpfxewendetswsgmys	86		
	1				
Qy	5	CYFNMAQYRGGEARFLPKDIDPSLCHLLLYAFAGMNHOLSTT-EWND-E-T-----LYQ	57		
Db	87	avtlrlctpnplkvllysgygnysaifcytlaskqterflksaalfiknmfdgdld	146		
	1				
Qy	58	EFNGLKKNMKNKLTLLAIGWMNGCTOKFTDMATANNRQTFVNSAIRFLRKYSFSDGLD	117		

```

Db 147 weyr-v-gv-a-e-eh-akveamktafeeaektsqkrl1llaasagkgtidsyve 200
118 WEPGSGSPAVDKEKFTTLVODLANAFQEAOTSGKERLLLSAAPAGOTYDAGYVD 177
QY 118 WEPGSGSPAVDKEKFTTLVODLANAFQEAOTSGKERLLLSAAPAGOTYDAGYVD 177
Db 201 slgknfdllfimsydlhgswekndlhgk1hptkgevsigifntefaadywaskmpke 260
178 KIQNNDPVLMAVDHSGSEKVTGHNSPLYKROESGAASLVDAVQOMLOKGPAS 237
QY 178 KIQNNDPVLMAVDHSGSEKVTGHNSPLYKROESGAASLVDAVQOMLOKGPAS 237
Db 261 kl1lg1pmvqgwtldnpeeta1gaasrpsasaknpagtsywe1ckylkegkety 320
238 KL1LGMPTTGRSTLSSSDTRVGAAPATGSGTGPPTKEGMLAYEVCSW-KGATKQRI 296
QY 238 KL1LGMPTTGRSTLSSSDTRVGAAPATGSGTGPPTKEGMLAYEVCSW-KGATKQRI 296
Db 321 hgevgaymwkgqgywydneetir1kmw1kkyggaftiawldfdftkscgkqpy 380
297 QDQKV-PYIFRDQWQWGFDDVESFKTKVSYLKQKGLGAMVWALDDDFAGFCSCNGRYP 355
QY 297 QDQKV-PYIFRDQWQWGFDDVESFKTKVSYLKQKGLGAMVWALDDDFAGFCSCNGRYP 355
Db 381 lnaissel 389
QY 356 LIQTLROEL 364

RESULT 10
ENTRY A56596 #type complete
TITLE chitinase (EC 3.2.1.14) - tobacco hornworm
ORGANISM #formal.name Manduca sexta #common.name tobacco hornworm
DATE 11-Aug-1995 #sequence.revision 11-Aug-1995 #text.change 03-May-1996
ACCESSIONS A56596
REFERENCE A56596
#authors Kramer, K.J.; Corpuz, L.; Choi, H.K.; Muthukrishnan, S.
#journal Insect Biochem. Mol. Biol. (1993) 23:691-701
#title Sequence of a cDNA and expression of the gene encoding epidermal and gut chitlinases of Manduca sexta.
#cross-references M0ID:93357793
#accession A56596
#status preliminary
#molecule_type mRNA
#residues 1-554 ##label KRA
#cross-references NCBI:136417; NCBI:136418
#experimental_source larvae
#note sequence extracted from NCBI backbone
KEYWORDS glycosidase; hydrolase
SUMMARY #length 554 #molecular-weight 62203 #checksum 4659

Query Match 37.4%; Score 1015; DB 12; Length 554;
Best Local Similarity 40.4%; Pred. No. 1,44e-160;
Matches 155; Conservative 98; Mismatches 107; Indels 24; Gaps 18;

```

```

QY 345 AGFSCNGRPLIQLTLOELSLPY 368
118 WEPGSGSPAVDKEKFTTLVODLANAFQEAOTSGKERLLLSAAPAGOTYDAGYVD 177
QY 118 WEPGSGSPAVDKEKFTTLVODLANAFQEAOTSGKERLLLSAAPAGOTYDAGYVD 177
Db 201 slgknfdllfimsydlhgswekndlhgk1hptkgevsigifntefaadywaskmpke 260
178 KIQNNDPVLMAVDHSGSEKVTGHNSPLYKROESGAASLVDAVQOMLOKGPAS 237
QY 178 KIQNNDPVLMAVDHSGSEKVTGHNSPLYKROESGAASLVDAVQOMLOKGPAS 237
Db 261 kl1lg1pmvqgwtldnpeeta1gaasrpsasaknpagtsywe1ckylkegkety 320
238 KL1LGMPTTGRSTLSSSDTRVGAAPATGSGTGPPTKEGMLAYEVCSW-KGATKQRI 296
QY 238 KL1LGMPTTGRSTLSSSDTRVGAAPATGSGTGPPTKEGMLAYEVCSW-KGATKQRI 296
Db 321 hgevgaymwkgqgywydneetir1kmw1kkyggaftiawldfdftkscgkqpy 380
297 QDQKV-PYIFRDQWQWGFDDVESFKTKVSYLKQKGLGAMVWALDDDFAGFCSCNGRYP 355
QY 297 QDQKV-PYIFRDQWQWGFDDVESFKTKVSYLKQKGLGAMVWALDDDFAGFCSCNGRYP 355
Db 381 lnaissel 389
QY 356 LIQTLROEL 364

RESULT 11
ENTRY A38368 #type complete
TITLE chitinase (EC 3.2.1.14) precursor - Bacillus circulans
ORGANISM #formal.name Bacillus circulans
DATE 28-Jun-1991 #sequence.revision 28-Jun-1991 #text.change 03-May-1996
ACCESSIONS A38368
REFERENCE A38368
#authors Watanabe, T.; Suzuki, K.; Oyanagi, W.; Ohnishi, K.; Tanaka, H.
#journal J. Biol. Chem. (1990) 265:15659-15665
#title Gene cloning of chitinase A1 from Bacillus circulans WL-12 revealed its evolutionary relationship to Serratia chitinase and to the type III homology units of fibronectin.
#cross-references M0ID:90368776
#accession A38368
#status preliminary
#molecule_type DNA
#residues 1-699 ##label WAT
#cross-references GB:J05599
CLASSIFICATION #superfamily fibronectin type III repeat homology
KEYWORDS glycosidase; hydrolase
SUMMARY #length 699 #molecular-weight 73677 #checksum 6757

Query Match 25.7%; Score 698; DB 10; Length 699;
Best Local Similarity 40.1%; Pred. No. 2,65e-101;
Matches 119; Conservative 73; Mismatches 85; Indels 20; Gaps 18;

```

```

Db 144 qnkl1qtnpnlk1tsvgwtw-smrfsdaataatrevfanasavdfrkynfdvld 202
58 EFNGLKKNPKLKTLLAIGWNGTQKFTDVAATANNROTGVNSAIRLRKYSFDLDD 117
QY 58 EFNGLKKNPKLKTLLAIGWNGTQKFTDVAATANNROTGVNSAIRLRKYSFDLDD 117
Db 203 wepvggldgsknpredqnyllsk1rek1daagavdk-kyl1t1asgasaya-a 260
118 WEPV-GS-QG-S-PAVDREFTTLVODLANAFQEAOTSGKERLLLSAAPAGOTYDAG 172
QY 118 WEPV-GS-QG-S-PAVDREFTTLVODLANAFQEAOTSGKERLLLSAAPAGOTYDAG 172
Db 261 ntelakiaivd1nmtvdfdgawg1sahnap1ydpdaasagypdantfvaagag 320
173 GYEVDAK1AONLDFVNMA1DFHGSWEKVTGHNSP-LYKQD-ESGNA-A-SLVDAVQO 228
QY 173 GYEVDAK1AONLDFVNMA1DFHGSWEKVTGHNSP-LYKQD-ESGNA-A-SLVDAVQO 228
Db 321 h1daipaak1lyvp1fyrgwd-gcagagn-gqyqctcgsavgtweasf1g1lean 378
229 W1QKGPASKL1GMPTGRSFTLASSDTRVGAAPATGSGTGPPTKEGMLAYEV-CS 287
QY 229 W1QKGPASKL1GMPTGRSFTLASSDTRVGAAPATGSGTGPPTKEGMLAYEV-CS 287
Db 379 y1kngy1tywndakvpy1lynsnkr1f1syddaesvyk1ay1ksg1lgan1fwe1 435
288 W-K-GATKQRIODQKPYIFRD-NQW-VGFDVESFKTKVSYLKQKGLGAMVWAL 339
QY 288 W-K-GATKQRIODQKPYIFRD-NQW-VGFDVESFKTKVSYLKQKGLGAMVWAL 339

RESULT 12
ENTRY S51591 #type complete
TITLE chitinase (EC 3.2.1.14)/lysozyme (EC 3.2.1.17) P2 precursor, pathogenesis-related - common tobacco
ORGANISM #formal.name Nicotiana tabacum #common.name common tobacco
DATE 15-Jul-1995 #sequence.revision 08-Sep-1995 #text.change 03-May-1996
ACCESSIONS S51591
REFERENCE S51591
#authors Heitz, T.; Second, S.; Kaufmann, S.; Geoffroy, P.; Prasad, V.; Brunner, F.; Fritlig, B.; Legend, M.
#journal Mol. Gen. Genet. (1994) 245:246-254
#title Molecular characterization of a novel tobacco pathogenesis-related (Pr) protein: a new plant chitinase/lysozyme.
#accession S51591
#molecule_type mRNA
#residues 1-378 ##label HEI
#cross-references EMBL:X78325

```

```

##experimental source cv. Samsun NN
#accession S51632
##molecule_type protein
##residues 3136;87-112;252-275;282-305;337-371 ##label HEW
CLASSIFICATION #superfamily Streptomyces chitinase ch140
KEYWORDS glycosidase; hydrolase
FEATURE
1-25 #domain signal sequence #status predicted #label SIG\
26-371 #product chitinase/lysozyme PZ #status predicted #label
371-378 MAT\
#domain carboxyl-terminal propeptide #status
SUMMARY #length 378 #molecular_weight 42019 #checksum 6689

Query Match 21.9%; Score 594; DB 5; Length 378;
Best Local Similarity 32.7%; Pred. No. 3,36e-82;
Matches 106; Conservative 87; Mismatches 108; Indels 23; Gaps 18;

43 nldstlfthfcafdpnsqnlispengds-fagfstvgkrnpvktfisiagqr 101
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
23 KDLDPSLCTHLIYAFAGMT--NHQLSTEWMDLRYGEFNG-LKKMPKLTILAT 78
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 102 a-dttaygimarqpnrsksfidssirlarfqfghlaldkeyp--1-s-aldtnlgill 156
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 79 NFGTKFMDNATANNRQTFVNSAIRFLRKYSFQGLDMDWEPESQSPAVDKRFRFTLV 138
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 157 newrtalmeernsgraillltaavsyprvnglnpyvesvarnlmnlmaydfygnw 216
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 139 QDLANAFQOEQLOSTKEREKLLSAVPAGQOTVYDGEYDKRDLAQNIDVNLMAVDFHGS-W 197
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 217 spsgnshnagfdpvnh--isgs--d9-anawtgaqvpckkllglpdygawrlvnpn 270
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 198 EKV-TGHSPLPKROESGAAASLNDNAVQMLQKGTSPKSLTLGMPYGRSFTLASS 256
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 271 lhdrlapaagksnva-vddgsm-tynlrlgylvsratvynaatlvgdyccysgmwisy 328
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 257 DTRVGAPATGCGTGPFPFKKEGMLAYEVCWKGATQR-IQDK-VP-YIFRDNQWVF 313
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 329 ddtgsvrnkvnyvkgrgllgyfaw 352
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 314 DVESFKTVSYLKGKLGAMWV 337
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 13
ENTRY S47133 #type complete
TITLE chitinase (EC 3.2.1.14) - fungus (Trichoderma harzianum)
ORGANISM #formal_name Trichoderma harzianum
DATE 23-Nov-1994 #sequence_revision 03-Aug-1995 #text_change 10-May-1996
ACCESSIONS S47133
REFERENCE S47133
AUTHORS Carasolio, C.; Gutierrez, A.; Jimenez, B.; van Montagu, M.;
Herrera-Estrella, A.
#submission submitted to the EMBL Data Library, June 1994
#description Characterization of ech-42, a Trichoderma harzianum
endochitinase gene expressed during mycoparasitism.
#accession S47133
#status preliminary
#molecule_type DNA
#residues 1-424 #label CAR
#cross-references EMBL:X79381

GENETICS
#introns 48/2; 81/2; 98/1
CLASSIFICATION #superfamily Streptomyces chitinase ch140
KEYWORDS glycosidase; hydrolase
SUMMARY #length 424 #molecular_weight 46298 #checksum 8773

Query Match 21.4%; Score 582; DB 5; Length 424;
Best Local Similarity 36.0%; Pred. No. 5.18e-80;
Matches 102; Conservative 69; Mismatches 95; Indels 17; Gaps 13;
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

QY	57	OEINGLKMMNKLTLLAIGMNGTOKFDMDVATANNROTFPVNSAIRFLRKISFSDCLD	116
Db	170	dweyp-addtqatcmvlllkeirsqdayaay-apy-yhllislaapagpehy-sfilhm	225
QY	117	DMEYPSGSGSPAVDKREFFTTLVODLANAFOEOAQTSCKERLLLSAAVPAQOTYVDAGY	176
Db	226	sdlqgvldvymlymadyagswsyysgldanlfanspmsnp-yntdqakdyikgvpva	284
QY	177	DKIAQNDLVNLMAYDHFSGWEKVTGHNSPYLRKQDESSGAASLMDVAAQOMQOKTPA	236
Db	285	skivlgmpliygraf-estqg-lyqcyvsqis-gsw-englwdy-kvlpkagatvyqd	336
QY	237	SKLLIGMPTYRSTTLASSSDSTRGAPATGSGTGPPTKGGMLATYEVCMSKCATQ--	294
Db	337	svagayysdpsskellisfdtpmdntkvsyiknlg1gsgmfw	379
QY	295	RIDOKVPYIFRDQNWGQFDDVESFRTKVSYLKQKGGAGAMW	337
RESULT	14		
ENTRY			
TITLE		JQI1975	#type complete
ORGANISM		chitinase (EC 3.2.1.14) 1 precursor - imperfect fungus (Aphanocladium album)	
DATE		03-Feb-1994	#sequence_revision 03-Feb-1994 #text_change 03-May-1996
ACCESSIONS		JQI1975	
REFERENCE		JQI1975	
#authors		Blaiseau, P.L.; Lafay, J.F.	
#journal		Gene (1992) 120:243-248	
#title		Primary structure of a chitinase-encoding gene (chl1) from the filamentous fungus Aphanocladium album: similarity to bacterial chitinases.	
#accession		JQI1975	
##molecule_type		mRNA	
##residues		1-423	##label BLA
##cross-references		GB:X61104	
COMMENT		This enzyme is essential for the degradation of insoluble chitin.	
GENETICS			
#gene		chl1	
CLASSIFICATION		47/2; 80/2; 97/1	
KEYWORDS		#superfamily Streptomyces chitinase chl40	
FEATURE		glycosidase; hydrolase	
1-34			
35-423		#domain signal sequence #status predicted #label SIG\	
SUMMARY		#product chitinase 1 #status predicted #label MAT	
		#length 423 #molecular_weight 46072 #checksum 8802	
Query Match		21.2%; Score 576; DB 5; Length 423;	
Best Local Similarity		36.2%; Pred. No. 6,42e-79;	
Matches 101; Conservative		65; Mismatches 95; Indels 18; Gaps 16;	
Db	115	Lkqnmkmkmlslygvtwst-nfpaassaatrkftgagsvfmkdwfgdididweyp	173
QY	62	LKKANPKLKITLLAGGNWFGQKTTDMVATANNQOTFVNSAIRFLRKYSFSDGLDMEYR	121
Db	174	-adatqagmmvlliyavrseldsyaay-ekg-ghfllislaapagdnhykikfa-e-1g	228
QY	122	GSQSPAVDKREFFTTLVODLANAFOEQAGQASGKERLLLSAAVPAQ-OTYVDAGYVDKIA	180
Db	229	kvldylnlmaydyagswsnytghdaanyaqpnpnatp-yntdawgavayingvpankiv	287
QY	181	QNLDFVNLMAYDHFSGWEKVTGHNSPYLRKQDESSGAASLMDVAAQOMQOKGTPASKLI	240
Db	288	lgmpliygrsfq-qt-eg-lykpyngis-gsw-englwdykalpk-agatvkcdtdak	339
QY	241	LGMPTYRSTTLASSSDSTRGAPATGSGTGPPTKGGMLATYEVCMSKCATQRIODCK	300
Db	340	gcysvdstekllisfdtpmstksvslkkojgsgmfw	378
QY	301	VPIYIF-RD-NQWGFDDVESFRTKVSYLKQKGGAGAMW	337

```

RESULT      15
ENTRY       S51369      #type complete
TITLE       Chitinase - fungus (Trichoderma harzianum)
ORGANISM    #formal_name Trichoderma harzianum
DATE        15-Jul-1995 #sequence_revision 19-Apr-1996 #text_change
            03-May-1996
ACCESSIONS  S51369
REFERENCE   S51369
#authors    Garcia, I.; Lora, J.M.; de la Cruz, J.; Benitez, T.; Llobell,
            A.; Pinfor-Toro, J.A.
#journal    Curr. Genet. (1994) 27:83-89
#title      Cloning and characterization of a chitinase (CHIT42) cDNA
            from the mycoparasitic fungus Trichoderma harzianum.
#accession  S51369
#status     Preliminary
#molecule_type DNA
#residues   1-423 #label GAR
SUMMARY     #length 423 #molecular_weight 46056 #checksum 2000

Query Match      20.6%; Score 561; DB 12; Length 423;
Best Local Similarity 35.7%; Pred. No. 3,43e-76;
Matches 101; Conservative 71; Mismatches 94; Indels 17; Gaps 13;

Db      111 qlfk-vkkanrglkvlslsgwtwt-nfpaastdanrknfaktaitfknkgfdgidi 168
QY      57 QEFNGLKKNPKLKTLLATIGMNFIOKFTDMVATANNROTENVSAIRFLRKISFDGLD 116

Db      169 dweyp-adatqaasm111kevrsgqdayaay-apy-yh111laapagkdy-sk1r1 224
QY      117 DWEXPGSGSPAVDKERFTLVQDLANAFQOEAGTSGKERLLLSAAVPAGOTYVDAGYEV 176

Db      225 adlqgvldylnlnaydyagsfsplctghdanlfnpsnpatp-fntdsavkdylingvpa 283
QY      177 DKIAQNDLDFVNLMAVDYFHGSMKEYTGHNSPLYKROESGAASLNVDAAYQOWLOKGTPA 236

Db      284 nkivlmpdygsf--qntag--lqctynvgs-gsw--eaqlwdyka1pk-agatvqyd 335
QY      237 SKLILGMPRTYGRSFTLASSDTRVGAPATSGTPGPTKEGMLAYTEVCWKGATKORI 296

Db      336 svakgyysynsackelstfdtpdminkvaylkslglgsmfw 378
QY      297 QDQKVPIYFRD--NQWVGFPDDESFKTKVSYLKGKLGAMW 337

```

Search completed: Fri Jun 27 17:06:56 1997  
 Job time : 134 secs.





\*\*\*\*\*  
 M E R E F  
 (TM)  
 \*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
 Distribution rights by IntelliGenetics, Inc.

Mparch\_p protein - protein database search, using Smith-Waterman algorithm  
 Run on: Fri Jun 27 17:07:13 1997; MasPar time 12.20 Seconds  
 Tabular output not generated. 648.449 Million cell updates/sec

Title: >US-08-663-618A-15  
 Description: (1-373) from US08663618A.pep  
 Perfect Score: 2717  
 Sequence: 1 AKLVCFYTNMAQYRGEARF.....YPLIQLTRQELSLPYLSSGT 373

Scoring table:  
 PAM 150  
 Gap 11

Searched: 59021 segs, 21210388 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: swiss-prot34  
 1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
 8:part8 9:part9 10:part10 11:part11

Statistics: Mean 48.971; Variance 85.460; scale 0.573

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1488	54.8	383	4	GP39_HUMAN CARTILAGE GLYCOPROTEI	0.00e+00
2	1086	40.0	504	2	CHIT_BRUVA ENDOCHITINASE PRECURS	1.03e-218
3	1015	37.4	554	2	CHIT_MANSE ENDOCHITINASE PRECURS	7.72e-202
4	698	25.7	699	2	CHIT_BACCT CHITINASE A1 PRECURSO	2.43e-127
5	576	21.2	423	2	CHIL_APHAL CHITINASE 1 PRECURSOR	3.13e-99
6	561	20.6	423	2	CHIL4_TRIHA 42 KD ENDOCHITINASE P	8.30e-96
7	546	20.1	427	2	CHIL_COCIM ENDOCHITINASE 1 PRECU	2.16e-92
8	437	16.1	563	2	CHIA_SERVA CHITINASE A PRECURSOR	7.14e-88
9	395	14.5	551	2	CHIT_NPVAC PROBABLE ENDOCHITINAS	1.33e-58
10	345	12.7	820	2	CHIA_ALTSO CHITINASE A PRECURSOR	9.66e-48
11	337	12.4	499	2	CHIB_SERVA CHITINASE B PRECURSOR	5.03e-46
12	301	11.1	619	2	CHIT_STRLI CHITINASE C PRECURSOR	2.19e-38
13	287	10.6	610	2	CHIT_STRPL CHITINASE 63 PRECURSO	1.86e-35
14	263	9.7	1146	5	KTXA_KILUA KILLER TOXIN ALPHA AN	1.67e-30
15	172	6.3	36	4	GP39_BOVIN CARTILAGE GLYCOPROTEI	8.53e-13
16	172	6.3	83	3	EDOS_PAPAN ESTRADIOL-DEPENDENT O	8.53e-13
17	121	4.5	597	2	CHIX_STRCI EXOCHITINASE 1 PRECUR	3.64e-04
18	108	4.0	452	4	HEMF_RHOSH COPROPORPHYRINOGEN II	3.11e-02
19	110	4.0	497	2	CHID_BACCI CHITINASE D PRECURSOR	1.60e-02
20	104	3.8	422	10	UMUC_ECOLI UMUC PROTEIN.	1.14e-01
21	100	3.7	364	11	YBR5_YEAST HYPOTHETICAL 40.9 KD	4.02e-01
22	101	3.7	1045	7	PRTT_SERMA EXTRACELLULAR SERINE	2.94e-01

RESULT ID	GP39_HUMAN STANDARD;	PRT;	383 AA.
AC	P36222; P30923;		
DT	01-JUL-1993 (REL. 26, CREATED)		
DT	01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)		
DT	01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)		
DE	CARTILAGE GLYCOPROTEIN-39 PRECURSOR (GP-39) (39 KD SYNOVIAL PROTEIN).		
OS	HOMO SAPIENS (HUMAN).		
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;		
OC	EUHARRIOTA; PRIMATES.		
RN	[1]		
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.		
RC	TISSUE-CARTILAGE;		
RX	MEDLINE: 9406458.		
RA	HAKALA B.E., WHITE C., RECKLIES A.D.;		
RL	J. BIOL. CHEM. 268:25803-25810(1993).		
RN	[2]		
RP	SEQUENCE OF 22-45.		
RA	MEDLINE: 90328983.		
RA	NYIRKOS P., GOLDS E.E.;		
RL	BIOCHEM. J. 269:265-268(1990).		
CC	- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CAPACITY OF CELLS TO		
CC	RESPOND TO AND COPE WITH CHANGES IN THEIR ENVIRONMENT.		
CC	- TISSUE SPECIFICITY: PRESENT IN ARTICULAR CHONDROCYTES, SYNOVIAL		
CC	CELLS AS WELL AS IN LIVER. UNDETECTABLE IN MUSCLE TISSUES, LUNG,		
CC	PANCREAS, MONONUCLEAR CELLS, OR FIBROBLASTS.		
CC	- PTM: GLYCOSYLATED.		
CC	- SUBCELLULAR LOCATION: EXTRACELLULAR.		
CC	- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.		
DR	EMBL: M80927; G348912; -		
DR	PIR: S10677; S10677.		
DR	PIR: A33162; A33162.		
DR	PROSITE: PS01095; CHITINASE_18.		
KW	GLYCOPROTEIN; SIGNAL.		
FT	SIGNAL 1 21		
FT	CHAIN 22 383		
FT	CARBOHYD 60 60		
FT	SEQUENCE 383 AA: 42613 KM; DDA5280F CMC32;		

Query Match 54.88; Score 1488; DB 4; Length 383;  
 Best Local Similarity 53.38; Pred. No. 0.00e+00;  
 Matches 195; Conservative 83; Mismatches 79; Indels 9; Gaps 5;

23 klvgyysvsgyregdscfpdldrlfclhlllyfsfnisnhdtkwkdvlygmnt 82  
 |||||:||||:||||:| | | |||||:||||:|:| | ||||| ||| :|  
 2 klvcyftnmaoyrgearfplkdldpslcthllyafagkgnholstewndetlyoeeng 61

```

Db      83 Lknrrpnllkllsvgwgnfsgqrfskiasntqsrtrfiksvpfllrhgfdglawlyp 142
Qy      62 LKRNKPKLKTLLIAGCNMFQCKQFTDMATVANNQSTVNSAIRLRYSPDGLDMEYP 121
Db      143 grf-----dkghftllikemkaefikeeq-pgkkgqllsaalsgkvtlssydlaktsg 166
Qy      122 GSQGSFPAVDKEFFTLVYDDLNNATQOEQKQKERRLLISAAPVPGQTYVPAEYVDKIAQ 181
Db      197 hldfslmtcyfghagwrgctgthgsplfrggedaspdfsrtdyavgymlrlqapasklvm 256
Qy      182 NDDFNFNMAIYDFHSGMKEVYTHNSPLKYRKQESSAASLNVDAVQOMLKGTPASKLLT 241
Db      257 giprfgrsflass-etvgaeplsgpypigrfikeagtlayaelcdrlrgatvhrtlrgq 315
Qy      242 GMPFRGSRFTLSSSDTRVGAAPATGSGTGPFBFTREGGLAVYEVCW-KGATQRIRIDQK 300
Db      316 vpyatkngqwygyddqesvkskvqylktrqlagamvvallddfqgsfcgqdlrfrpltna 375
Qy      301 VPYIRDNQWNGVGFDEVEFSKTKVSYLKOKGGLGAMVVALDLDDPAGFSQNG-NYPLIQF 359
Db      376 lkdaa 381
Qy      360 LKQELUS 365

```

RESULT ID	2	CHIT. BRDUA	STANDARD:	PRT:	504 AA.
AC	P23030.				
DT	01-DEC-1992	(REL. 24, CREATED)			
DT	01-DEC-1992	(REL. 24, LAST SEQUENCE UPDATE)			
DT	01-FEB-1995	(REL. 31, LAST ANNOTATION UPDATE)			
DE	ENDOCHITINASE	PRECURSOR (EC 3.2.1.14) (MEL ANTIGEN).			
OS	BRIGIA MALAYI.				
OC	EURAROTEA, METAZOA; ACCELLOMATES; NEMATODA; SECCERINENTEA; SPIRURIDA.				
RN	[1]				
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.				
RA	MEDLINE; 921179220.				
RA	FURMAN J.A., LANE W.S., SMITH R.F., PIESSENS W.F., PERLER F.B.;				
RL	PROC. NATL. ACAD. SCI. U.S.A. 89:1548-1552(1992).				
CC	-1- FUNCTION: THE MEL ANTIGEN IS A MICROFILARIAE CHITINASE, WHICH MAY				
CC	-1- FUNCTION TO DEGRADE CHITIN-CONTAINING STRUCTURES IN THE MICRO-				
CC	-1- FILARIA OR IN ITS MOSQUITO VECTOR DURING PARASITE DEVELOPMENT AND				
CC	-1- TRANSMISSION.				
CC	-1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE BETA-1,4 LINKAGES OF				
CC	-1- N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.				
CC	-1- PTM: O-GLYCOSYLATED.				
CC	-1- KNOWN TO BIND CALCIUM.				
CC	-1- DEVELOPMENTAL STAGE: THE APPEARANCE OF THE MEL ANTIGEN CORRESPOND				
CC	WITH THE ONSET OF THE PARASITE'S ABILITY TO INFECT THE MOSQUITO.				
CC	-1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL				
CC	-1- HYDROLASES).				
DR	EMBL; M73689; G156064; -.				
DR	PIR; A38221; A38221.				
DR	PROSITE; PS01095; CHITINASE_18.				
KM	HYDROLASE; GLYCOSIDASE; CHITIN DEGRADATION; SIGNAL; ANTIGEN;				
KM	REPEAT; GLYCOPROTEIN; CALCIUM-BINDING.				
KW	REPEAT				
FT	SIGNAL	1 22			
FT	CHAIN	23 504	ENDOCHITINASE.		
FT	DOMAIN	23 400	CATALYTIC.		
FT	DOMAIN	401 450	SER/THR-RICH (LINKER).		
FT	DOMAIN	407 448	3 X 14 AA APPROXIMATE TANDEM REPEATS.		
FT	ACT_SITE	148 148	PROTON DONOR (BY SIMILARITY).		
SO	SEQUENCE	504 AA; 55971 MM; 404755B CRC32;			

Query Match	40.0%;	Score 1086;	DB 2;	Length 504;
Best Local Similarity	42.8%;	Pred. No. 1.03e-218;		
Matches 156; Conservative		94;	Mismatches 102;	Indels 15; Gaps 11;

Db 27 cyttwagrydgsegkfllpgnlpngctihllyafakvdelgdsfpfwendedeletwskgmys 86  
|||:|||||:::||||:||||:| ||||| :|  
QY 5 CYFTNNAOYRQGARFLPRKDLDPISLCTHLIYAFAGMTNHQLSTT-ENWDE-T-----LYQ 57

```

Db 87 avtklirempojlxvllsyggyngfsaalfcylaksaqktefiflsalaflknmfdofdid 146
Qy 58 EFNGIKKXNPKLTKTLAIGGNEFGTOFTVMVAANNRQTFVNSAIRFLFKRYSEFDGDI 117
Db 147 weyp-v-9v-a-e-eh-aklveamktafveeaktsyqrrlllaavysaggtldsygne 200
Qy 118 WEYPOSGSPVADWERFTTLVQDANNAFOQAQTSIGKERLLLSNAVPAGGTYVDAGIEVD 177
Db 201 slgknfdllfmsydlngswekenvdlhgkllhpckevsygitinfetaadyaskmpe 260
Qy 178 KIAQNLDPVNNALMDFHGSMWKYGHNSPLYKRQEEGGAASLNVDAVOQMLOKGPAS 237
Db 261 klllgipmyagqwtclnmpetaiaaasrpsaaklnpagtasyewlckylkeggkety 320
Qy 238 KLILCMPTYGRSFTLASSDDTRVOCAPATGSGTPEPFRKKEGMLAYVECSW-KGATQR 296
Db 321 hgeyvgaymvksgdyvygdyneetlrlkmkwlkekyggaafiwaldfdffgskscgkypyp 380
Qy 297 ODOKV-PIYFRNDNOMVGGDDVESEFKRVSYLKQKGLGAVWALDDDFEFGSCNOGRYP 355
Db 381 llnaisel 389
Qy 356 LIQTLRQEL 364

```

RESULT	3	STANDARD:	PRT:	554 AA.
ID	CHIT_MANSE			
AC	P36362;			
DT	01-JUN-1994 (REL. 29, CREATED)			
DT	01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)			
DT	01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)			
DE	ENDOCHITINASE PRECURSOR (EC 3.2.1.14)			
OS	MANDUCA SEXTA (TOBACCO HAWMOTH) (TOBACCO HORRWORM).			
OC	EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; LEPIDOPTERA.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE: 93357793.			
RA	KRAMER K.J., CORPUZ L., CHOI H.R., MUTHUKRISHNAN S.;			
RL	INSECT BIOCHEM. MOL. BIOL. 23:691-701(1993).			
CC	-1- FUNCTION: DIGEST CHITIN IN THE EXOSKELETON DURING THE MOLTING PROCESS.			
CC	-1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE BETA-1,4 LINKAGES OF			
CC	N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.			
CC	-1- DEVELOPMENTAL STAGE: HIGH LEVELS SEEN IN THE EPIDERMIS ON DAY 0,			
CC	BUT RAPIDLY DISAPPEARS AND IS UNDETECTED ON DAYS 1-4 OF FIFTH			
CC	INSTAR. IT REAPPEARS ON DAY 5 AND PEAKS ON DAY 7 AFTER WHICH A			
CC	RAPID DECLINE IS SEEN. IN THE GUT IS DETECTED ON DAY 6 WITH LOWER			
CC	LEVELS SEEN ON DAYS 0, 7 AND 8.			
CC	-1- TISSUE SPECIFICITY: EPIDERMIS AND GUT.			
CC	-1- SUBCELLULAR LOCATION: SECRETED.			
CC	-1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL			
CC	HYDROLASES).			
DR	EMBL: U02270; G406049; -;			
DR	N-ACETYL-564757; -; NOT_ANNOTATED_CDS.			
DR	HSSP: P07254; ICTN.			
DR	PROSITE: PS01095; CHITINASE.18.			
KW	HYDROLASE; GLYCOSIDASE; CHITIN DEGRADATION; SIGNAL; GLYCOPROTEIN.			
FT	SIGNAL	1	19	POTENTIAL.
FT	CHAIN	20	554	ENDOCHITINASE.
FT	DOMAIN	396	453	SER/THR-RICH.
FT	ACT_SITE	146	146	PROTON DONOR (BY SIMILARITY).
FT	CARBOHYD	85	85	POTENTIAL.
FT	CARBOHYD	303	303	POTENTIAL.
FT	CARBOHYD	407	407	POTENTIAL.
FT	CARBOHYD	545	545	POTENTIAL.
SO	SEQUENCE	554 AA;	62203 MW;	FA8/F8AD CRC32;

Query Match	37.48;	Score 1015;	DB 2;	Length 554;
-------------	--------	-------------	-------	-------------

Matches 155; Conservative 98; Mismatches 107; Indels 24; Gaps 18;

```

Db 23 arivcyfsmavvypgvrgyriedipvekcthllysfivgvegnsevlldpeldvdkng 82
   ||::|||::||| || |::| |::|||::|||::|::|::|::|::|::|::|::|::|

```

Matches	119;	Conservative	73;	Mismatches	85;	Indels	20;	Gaps	18
Db	144	qlnkjktgnlpnkllisvgvwtw-snifsdvaataatrevfanaavgdflrkynfdgvdld	202						
Oy	58	EFNGIKKMPKLTLLAIGWNGFTQKFTDVAATANNQTFVNSAIRFLRKXSFEDGLDLD	117						
Db	203	weyrpsggldgnskrpdkqnytllektrkidaaavdgk-kyllttiasgaaty-a	260						
Oy	118	WEYR-GS-QG-S-PAYDKERFTLVODLANAFQOEQTSKREKRLLSAAVPAQOTVDA	172						
Db	261	ntelakiaiatdwimimtydfngawqklsanahplnydpasaaavgdydanfvaagag	320						
Oy	173	GVEVDKIQONIDFYNALAYDFHSGMEKVTGNNSFL-YKRQE-ESGAA-A-SLNDAAVQO	228						
Db	321	hldavgpaaklvlgvpfgrgdw-gcaqaagn-ggyqctcgssvgtweagafdfydean	378						
Oy	229	WLQKSTPASKILLGMPRTYGRSFTLASSSDRDVGAAPATGSGTPGFTREGCLATYEV-CS	287						
Db	379	ylnkngyltryndactakvpylylnasnkrffisyddaeavgktyaklskylggamfwel	435						
Oy	288	W-K-GAKRKRIODQKVPYIFRD-NQM-VGDDVSEFTKYSYKQGLGGAMWAL	339						
RESULT	5								
ID	CHIL_APHAL	STANDARD;	PRT;	423	AA.				
AC	P32470;								
DT	01-OCT-1993	(REL. 27, CREATED)							
DT	01-FEB-1996	(REL. 33, LAST SEQUENCE UPDATE)							
DT	01-FEB-1996	(REL. 33, LAST ANNOTATION UPDATE)							
DE	CHITINASE 1	PRECURSOR (EC 3.2.1.14).							
GN	CHIL.								
OS	APHANOCLOIDIUM ALBUM.								
OC	EUBARYOTA; FUNGI; DEUTEROMYCOTINA (IMPERFECT FUNGI).								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN-ETHM 483;								
RA	BLAISEAU P.-U., LAFAI J.-F.;								
RL	GENE 120:243-248(1992).								
RN	[2]								
RP	SEQUENCE OF 35-57.								
RC	MEDLINE; 92136437.								
RA	BLAISEAU P.-U., GRISON R., BERTHEAU Y., BRYGOO Y.;								
RL	CURR. GENET. 21:61-66(1992).								
CC	-1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE BETA-1,4 LINKAGES OF								
CC	-1- N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.								
CC	-1- SUBCELLULAR LOCATION: SECRETED.								
CC	-1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL								
CC	HYDROLASES).								
DR	EMBL; X64104; G429026; -.								
DR	PIR; J01975; J01975.								
DR	HSSP; P07254; ICTN.								
DR	PROSITE; PS01095; CHITINASE_18.								
DR	HYDROLASE; GLYCOSIDASE; CHITIN DEGRADATION; SIGNAL; ZYMOGEN.								
FT	SIGNAL	1	22	POTENTIAL.					
FT	PROPEP	23	34	POTENTIAL.					
FT	CHAIN	35	423	CHITINASE 1.					
FT	ACT_SITE	171	171	PROTON DONOR (BY SIMILARITY).					
SO	SEQUENCE	423	AA;	46058	MM;	884566E6	CRC32;		
Query Match		21.2%;	Score 576;	DB 2;	Length 423;				
Best Local Similarity		36.2%;	Pred. No. 3.13e-99;						
Matches	101;	Conservative	65;	Mismatches	95;	Indels	18;	Gaps	16;
Db	115	lkkqrmkmvmlsigvwtw-nfpaasaasatrkftagsavgfmkdwfgldidweyp	173						
Oy	62	LKKMMPKLTLLAIGWNGFTQKFTDVAATANNQTFVNSAIRFLRKXSFEDGLDWEYP	121						
Db	174	-adatagqnmvllqavrseldsyaay-akg-nhflfslsaapgdghnykklka-e-lg	228						
Oy	122	GSQGSPPADKERRFTTLVODLANAFQOEQTSYSGKERLLLSAIVPAAG-QTYVDAGEYVKA	180						
Db	229	kvlidinlmaydygswsnfyqbhanlyanqpnnatcp-yntddavayinqvpankiv	287						

QY 181 QNLEFVNLMAYDFHGSWEKVTGHNNSPLYKROESGAASLVDAVQOOLKGTPTASKLI 240  
Db 288 lgmplgyrsf-qf--eg-igkpyngis-gsw--englwykalkp-agnetkcdclak 339  
QY 241 LGMPYGRSFTLASSDSDRVGAPATGSGTPGFTKEGMLAYEVCSWKATKORIODOK 300  
Db 340 gcsydpstkelisfdtpamistkxwjkqkylgntmf 378  
QY 301 VPYIF-RD-NQWGFDDVESFKTKVSYLKOKGLGAMW 337

RESULT 6  
ID CH14-TRHA STANDARD; PRT; 423 AA.  
AC P48827;  
DT 01-FEB-1996 (REL. 33, CREATED)  
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
DE 42 KD ENDOCHITINASE PRECURSOR (EC 3.2.1.14).  
GN CHIT42.  
OS TRICHODERMA HARZIANUM.  
CC EUKARYOTA; FUNGI; DEUTEROMYCOTINA (IMPERFECT FUNGI).  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 35-52; 93-107; 371-385 & 397-414.  
RX MEDLINE; 95269313.  
RA GARCIA I., LORA J.M., LA CRUZ J., BENITEZ T., LLOBELL A.,  
RA PINTOR-TORO J.A.;  
RL CURR. GENET. 27:83-89(1994).  
CC -1- FUNCTION: MORPHOGENETIC ROLE DURING APICAL GROWTH, CELL DIVISION  
AND DIFFERENTIATION (CELL WALL MORPHOGENESIS). ANTIFUNGAL  
AGENT.  
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE BETA-1,4 LINKAGES OF  
N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.  
CC -1- SUBCELLULAR LOCATION: SECRETED.  
CC -1- INDUCTION: SPECIFICALLY INDUCED BY CHITIN AND IS CARABOLITE  
REPRESSIBLE.  
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
HYDROLASES).  
EMBL: S78423; G999376; -;  
DR HYDROLASE; GLYCOSIDASE; CHITIN DEGRADATION; SIGNAL; ZMOGEN;  
KW CHITIN-BINDING.  
FT SIGNAL 1 22 POTENTIAL.  
FT PROPEP 23 34  
FT CHAIN 35 423 42 KD ENDOCHITINASE.  
FT ACT SITE 171 171 PROTON DONOR (BY SIMILARITY).  
FT CARBOHYD 218 218  
SQ SEQUENCE 423 AA; 46056 MW; 899DA50A CRC32;

Query Match 20.6%; Score 561; DB 2; Length 423;  
Best Local Similarity 35.7%; Pred. No. 8.30e-96;  
Matches 101; Conservative 71; Mismatches 94; Indels 17; Gaps 13;

Db 111 qlfk-vkkaanglkvlisigwtwst-nfpaastdnrknfakalcftmkwfgdgidl 168  
QY 57 QEFNLKTKMNPFLKTLILAGWNFGTOKFTDMVATANNKQTFVNSAIRLKRYSFDGLDL 116  
Db 169 dweyp-adatqasnmlllkewrsgrdayaay-epg-yhflitlaapagkny-skiri 224  
QY 117 DWEYFGSGSPVDEKERTTLVQDILANAFQQAQTSKGERLLLSAIVPAQOTYVDAGYEV 176  
Db 225 adlqgvdylinmaydyagsfsglqhdanlfnpnnapcp-fatdaavkdyingvpa 283  
QY 177 DKINQNLDFVNLMAYDFHGSWEKVTGHNNSPLYKROESGAASLVDAVQOOLKGTPTA 236  
Db 284 nkivlgmplgyrsf--qntag--igklyngvgs-gsw--eaglwkykalkp-agnetvqd 335  
QY 237 SKLITGMPYGRSFTLASSDSDRVGAPATGSGTPGFTKEGMLAYEVCSWKATKORI 296  
Db 336 svakgyysnatkelisfdtpdmintkvaaylksglgsmtw 378  
QY 297 ODOKVPIFRD--NQWGFDDVESFKTKVSYLKOKGLGAMW 337

RESULT 7  
ID CH11-COCIN STANDARD; PRT; 427 AA.  
AC P54196;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE ENDOCHITINASE 1 PRECURSOR (EC 3.2.1.14) (COMPLEMENT-FIXATION ANTIGEN)  
DE (CF-ANTIGEN) (CF-AG).  
GN CTS1.  
OS COCCIDIODES IMMITIS.  
CC EUKARYOTA; FUNGI; DEUTEROMYCOTINA (IMPERFECT FUNGI).  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-C735;  
RX MEDLINE; 96144270.  
RA PISHKO E.J., KIRKLAND T.N., COLE G.T.;  
RL GENE 167:173-177(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SILVEIRA;  
RA YANG C., ZHU Y., MAGEE D.M., COX R.A.;  
RL SUBMITTED (MAR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE BETA-1,4 LINKAGES OF  
N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.  
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
HYDROLASES).  
EMBL: L41663; G1200190; -;  
DR EMBL; U51271; G1256769; -;  
DR EMBL; U33265; G1255728; -;  
KW HYDROLASE; GLYCOSIDASE; CHITIN DEGRADATION; CHITIN-BINDING; SIGNAL;  
KW GLYCOPROTEIN.  
FT SIGNAL 1 2 POTENTIAL.  
FT CHAIN 2 427 ENDOCHITINASE 1.  
FT CARBOHYD 387 387  
FT CONFLICT 15 47  
FT CONFLICT 15 47  
FT CONFLICT 15 47  
FT CONFLICT 15 47  
SQ SEQUENCE 427 AA; 47629 MW; 6CB9AF73 CRC32;

Query Match 20.1%; Score 546; DB 2; Length 427;  
Best Local Similarity 34.1%; Pred. No. 2.16e-92;  
Matches 95; Conservative 71; Mismatches 94; Indels 19; Gaps 17;

Db 115 lknmrnlkllisigwty-spntktpasteegkktadsiklmkldgfdgididweyp 173  
QY 62 LKNMNPFLKTLILAGWNFGTOKFTDMVATANNKQTFVNSAIRLKRYSFDGLDLDEYF 121  
Db 174 ede-kqandfyllkacrealdaysak-hpnqk-kfllitlaspaqpnq-nk-lklaemd 228  
QY 122 GSQSPAVDKRFRFTTLVQDILANAFQQAQTSKGERLLLSAIVPAQ-QTYVDAGYEVDKIA 180  
Db 229 kylfdwlmaydfsgswdkvsgvmsnvfpstckp-estpfisdkaavkdylkaypankiv 287  
QY 181 QNLEFVNLMAYDFHGSWEKVTGHNNSPLYKROESGAASLVDAVQOOLKGTPTASKLI 240  
Db 288 lgmplgyrsf--ast-dg-lgtfingvgs-gsw--engvwydkmpq-qgsqvleldia 339  
QY 241 LGMPYGRSFTLASSDSDRVGAPATGSGTPGFTKEGMLAYEVCSWKATKORIODOK 300  
Db 340 asysydknryliysdytkvklagkaeyltkngmvggm-w 377  
QY 301 VPYIF-RDQW-VGPFDDVESFKTKVSYLKOKGLGAMW 337  
RESULT 8  
ID CH1A-SERMA STANDARD; PRT; 563 AA.  
AC P07254;  
DT 01-APR-1988 (REL. 07, CREATED)  
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)  
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
DE CHITINASE A PRECURSOR (EC 3.2.1.14).  
GN CH1A.  
OS SERRATIA MARCESCENS.

OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
 OC ENTEROBACTERIACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA KOO J.C., LIM C.O., CHOI Y.J., KIM C.Y., BARK J.D., LEE S.Y.,  
 RA CHO M.J.;  
 RL SUBMITTED (JAN-1993) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 990 / QMB1466;  
 RA JONES J.D.G., GRADY K.L., SUSLOW T.V., BEDBROOK J.R.;  
 RL EMBO J. 5:467-473(1986).  
 RN [3]  
 RP REVISIONS, AND X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
 RX MEDLINE: 95219379.  
 RA PERRAKIS A., TEWS I., DAUTER Z., OPPENHEIM A.B., CHET I., WILSON K.S.,  
 RA VORGTAS C.E.;  
 RL STRUCTURE 2:1169-1180(1994).  
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE BETA-1,4 LINKAGES OF  
 CC N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.  
 CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
 CC HYDROLASES).  
 DR EMBL: L01455; G152818; -  
 DR EMBL: X03657; G46831; -  
 DR PIR: A25090; A25090.  
 DR HSSP: P07254; 1CTN.  
 DR PROSITE: PS01095; CHITINASE\_18.  
 KM HYDROLASE; GLYCOSIDASE; CHITIN DEGRADATION; SIGNAL.  
 FT SIGNAL 1 23  
 FT CHAIN 24 563  
 FT DOMAIN 150 563  
 FT ACT\_SITE 315 315  
 FT ACT\_SITE 391 391  
 FT ACT\_SITE 391 391  
 FT CONFLICT 76 77  
 FT CONFLICT 139 139  
 FT CONFLICT 226 226  
 FT CONFLICT 395 395  
 FT CONFLICT 410 430  
 FT CONFLICT 437 437  
 FT CONFLICT 464 467  
 FT CONFLICT 473 473  
 FT CONFLICT 484 484  
 SQ SEQUENCE 563 AA: 60979 MW: 7764DB57 CRC32;  
 Query Match 16.1%; Score 437; DB 2; Length 563;  
 Best Local Similarity 28.8%; Pred. No. 7.14e-68;  
 Matches 93; Conservative 84; Mismatches 123; Indels 23; Gaps 19;

RESULT 9  
 ID CHIT\_NPYAC STANDARD; PRT: 551 AA.  
 AC P41684;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE PROBABLE ENDOCHITINASE PRECURSOR (EC 3.2.1.14).  
 OS AUTOGRAHA CALIFORNICA NUCLEAR POLYEDROSIS VIRUS (ACANPV).  
 OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; BACULOVIRIDAE; EUBACULOVIRINAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-C6;  
 RX MEDLINE: 94303173.  
 RA AYRES M.D., HOWARD S.C., KUZIO J., LOPEZ-FERRER M., POSSEE R.D.;  
 RL VIROLOGY 202:586-605(1994).  
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE BETA-1,4 LINKAGES OF  
 CC N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.  
 CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
 CC HYDROLASES).  
 DR EMBL: L22858; G559195; -  
 DR PROSITE: PS01014; ER\_TARGET.  
 DR PROSITE: PS01095; CHITINASE\_18.  
 KM HYDROLASE; GLYCOSIDASE; CHITIN DEGRADATION; SIGNAL; GLYCOPROTEIN.  
 FT SIGNAL 1 17  
 FT CHAIN 18 551  
 FT ACT\_SITE 305 305  
 FT CARBOHYD 173 173  
 FT CARBOHYD 444 444  
 FT SITE 548 551  
 FT SITE 548 551  
 SQ SEQUENCE 551 AA: 61368 MW: BDB124DB CRC32;  
 Query Match 14.5%; Score 395; DB 2; Length 551;  
 Best Local Similarity 29.2%; Pred. No. 1.33e-58;  
 Matches 95; Conservative 80; Mismatches 122; Indels 28; Gaps 27;

Db 231 pfaalqagkyavawddpykgnfsgdlnakqahpdlkllpslsgwtlsdpffmgdkvk 289  
 Oy 36 AFAGMTNHQSTWENDE-T-LYOEFNGLKMPKLTLLAIGMNGTGCTFTDMVATAN 93  
 Db 290 -rtrfsgvvefqtgkffgvdldwefpgkganplsgpdegtvllmkellmldq 348  
 Oy 94 NRQTFVNSAIRFLKYSF-DGLDDMEYPSQGS-PAV---DKERTTLVQDANAFQO 147  
 Db 349 lsaetgr-kyeltsalsagdkldkavynag--nsmdhlfmsydfypfdlknqht 405  
 Oy 148 EAQTSKGERLLLSAAVPAQGTVD-AGEYVDKLAQNLDVNLMAVDHSGWE-KVIGHNS 205  
 Db 406 alnapavkptayt-tvng-vnaallagvkvkvvgvtamyrgrvgtvngynqnlpftgt 463  
 Oy 206 PLYKROESGAASLNDVAVOQMLQKTPASKLLILGMPYGRSFTLASSSDTRVGPAT 265  
 Db 464 atg-pvrgtvgknjivdrgagqfmsgwytydataeayvfvfpgsdlitddasvq 522  
 Oy 266 GSGPGPFTEGGLAYEVCS-WKGATKO-RIOD-QKVPYFR-DN-QWVGFDVESFK 320  
 Db 523 aksgyvldkqglgfweladn 545  
 Oy 321 TKVSYLKGKGLGAMVWALDD 343

RESULT 10  
 ID CHIA\_ALRNO STANDARD; PRT: 820 AA.  
 AC P32823;  
 DT 01-OCT-1993 (REL. 27, CREATED)  
 DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
 DE CHITINASE A PRECURSOR (EC 3.2.1.14) (CHI-A).  
 GN CHIA.  
 OS ALTERNOMAS SP. (STRAIN O-7).  
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
 RN [1]





